

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2003, 23:34:11 : Search time 1733 Seconds  
(without alignments)  
8073.337 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 342  
Sequence: 1 GCAGAGCTGAGGGGATGTGG.....CTTCAGTAAATATGTACA 342

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank:

- 1: gb\_ba.\*
- 2: gb\_htg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vl.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_htg\_hum.\*
- 31: em\_htg\_inv.\*
- 32: em\_htg\_other.\*
- 33: em\_htg\_mus.\*
- 34: em\_htg\_pln.\*
- 35: em\_htg\_rod.\*
- 36: em\_htg\_man.\*
- 37: em\_htg\_vrt.\*
- 38: em\_sy.\*
- 39: em\_htgo\_hum.\*
- 40: em\_htgo\_mus.\*
- 41: em\_htgo\_other.\*

Pred. No. Is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	342	100.0	420	6	AR110201	AR110201 Sequence
2	342	100.0	480	6	AR110202	AR110202 Sequence
3	342	100.0	615	6	AR071301	AR071301 Sequence
4	342	100.0	617	9	HUMINSH	L34838 Homo sapien
5	342	100.0	618	6	A47856	A47856 Sequence 1
6	342	100.0	618	6	A98978	A98978 Sequence 1
7	342	100.0	618	6	AX019983	AX019983 Sequence
8	342	100.0	618	6	BD107918	BD107918 Expressio
9	342	100.0	673	9	BC026254	BC026254 Homo sapi
10	342	100.0	934	6	AR126781	AR126781 Sequence
11	342	100.0	934	6	AR202438	AR202438 Sequence
12	340.4	99.5	615	6	AR071299	AR071299 Sequence
13	340.4	99.5	615	6	AR126780	AR126780 Sequence
14	340.4	99.5	615	6	AR202437	AR202437 Sequence
15	224.2	70.5	482	11	G27877	G27877 human STS
16	224.4	65.6	108265	9	AL133547	AL133547 Human DNA
17	153	44.7	153	6	AR126784	AR126784 Sequence
18	153	44.7	153	6	AR202441	AR202441 Sequence
19	115.6	33.8	108265	9	AL133547	AL133547 Human DNA
20	99	28.9	123	6	AR126783	AR126783 Sequence
21	99	28.9	123	6	AR202440	AR202440 Sequence
22	90	26.3	93	6	AR126785	AR126785 Sequence
23	90	26.3	93	6	AR202442	AR202442 Sequence
24	88	25.7	801	4	AB000201	AB000201 Equus cab
25	86.4	25.3	428	4	S78800	S78800 relaxin B,C
26	84.8	24.8	549	6	A06852	A06852 S.scrofa mR
27	84.8	24.8	549	6	A17335	A17335 S.scrofa mR
28	84.8	24.8	549	6	I02482	I02482 Sequence 3
29	84.8	24.8	733	4	P1GRELX	K01088 Pig relaxin
30	84.8	24.8	752	6	A06051	A06051 Recombinant
31	84.8	24.8	752	6	A16593	A16593 S.scrofa mR
32	84.8	24.8	753	6	E00122	E00122 mRNA coding
33	79.6	23.3	600	4	AF254739	AF254739 Camelus d
34	78	22.8	543	4	AF233688	AF233688 Felis cat
35	73.4	21.5	567	6	I05447	I05447 Sequence 3
36	71.8	21.0	312	6	I00753	I00753 Sequence 13
37	71.8	21.0	501	9	P1RLX4	Z27245 P.trogiodi
38	71.8	21.0	558	6	A08925	A08925 H.sapiens m
39	71.8	21.0	558	6	A17315	A17315 H.sapiens m
40	71.8	21.0	558	6	E00219	E00219 cDNA encodi
41	71.8	21.0	558	6	I00742	I00742 Sequence 2
42	71.8	21.0	558	6	I00749	I00749 Sequence 9
43	71.8	21.0	558	9	HSREL3	X00948 Human mRNA
44	71	20.8	156592	9	AL135786	AL135786 Human DNA
45	70.2	20.5	558	6	A06919	A06919 H.sapiens m

ALIGNMENTS

RESULT 1  
AR110201  
LOCUS AR110201  
DEFINITION Sequence 1 from patent US 6114307.  
ACCESSION AR110201  
VERSION AR110201.1 GI:12826477  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 420)  
AUTHORS Jaspers,S.R., Sprugel,K.H., Ren,H.Ping., Humes,J.M. and Conklin,D.C.  
TITLE Methods for stimulating pancreatic islet cell regeneration  
JOURNAL Patent: US 6114307-A 1 05-SEP-2000;

```
FEATURES          Location/Qualifiers
Source            1..420
BASE COUNT       124 a 100 c 98 g 98 t
ORIGIN
Query Match      100.0%; Score 342; DB 6; Length 420;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAAACACTTGTGTCATATATGCCCCATG 60
Db 76 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAAACACTTGTGTCATATATGCCCCATG 135
Qy 61 CCTGAGNAGACATTCACCCACCACCCAGAGGGTGGCTGCTGATCTGGAGTCCCAA 120
Db 136 CCTGAGNAGACATTCACCCACCACCCAGAGGGTGGCTGCTGATCTGGAGTCCCAA 195
Qy 121 GAAATGGTGTCAACCTCCCAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180
Db 196 GAAATGGTGTCAACCTCCCAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 255
Qy 181 ATTCCTAATTTGTACAGAGCTGGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 240
Db 256 ATTCCTAATTTGTACAGAGCTGGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 315
Qy 241 AATATATATCTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATCTGTTGT 300
Db 316 AATATATATCTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATCTGTTGT 375
Qy 301 GAAATATATCTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATCTGTTGT 342
Db 376 GAAATATATCTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATCTGTTGT 417

RESULT 2
AR110202
LOCUS          AR110202          480 bp      DNA      linear      PAT 14-FEB-2001
DEFINITION     Sequence 3 from patent US 6114307.
ACCESSION      AR110202
VERSION        AR110202.1 GI:12826478
KEYWORDS       .
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 480)
AUTHORS        Jaspers S.R., Sprugel K.H., Ren H.Ping., Humes J.M. and
                Conklin D.C.
TITLE          Methods for stimulating pancreatic islet cell regeneration
JOURNAL        Patent: US 6114307-A 3 05-SEP-2000;
FEATURES       Location/Qualifiers
Source         1..480
BASE COUNT     136 a 114 c 121 g 109 t
ORIGIN
Query Match    100.0%; Score 342; DB 6; Length 480;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAAACACTTGTGTCATATATGCCCCATG 60
Db 136 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAAACACTTGTGTCATATATGCCCCATG 195
Qy 61 CCTGAGNAGACATTCACCCACCACCCAGAGGGTGGCTGCTGAAATCTGGAGTCCCAA 120
Db 196 CCTGAGNAGACATTCACCCACCACCCAGAGGGTGGCTGCTGAAATCTGGAGTCCCAA 255
Qy 121 GAAATGGTGTCAACCTCCCAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180
Db 256 GAAATGGTGTCAACCTCCCAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 315
Qy 181 ATTCCTAATTTGTACAGAGCTGGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 240
```

```
Db 316 ATTCCTAATTTGTACAGAGCTGGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 375
Qy 241 AAAATAATATCTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATCTGTTGT 300
Db 376 AAAATAATATCTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATCTGTTGT 435
Qy 301 GAAATATATCTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATCTGTTGT 342
Db 436 GAAATATATCTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATCTGTTGT 477

RESULT 3
AR071301
LOCUS          AR071301          615 bp      DNA      linear      PAT 18-FEB-2000
DEFINITION     Sequence 3 from patent US 5910480.
ACCESSION      AR071301
VERSION        AR071301.1 GI:7222189
KEYWORDS       .
SOURCE         Unknown.
ORGANISM       Unclassified.
REFERENCE      1 (bases 1 to 615)
AUTHORS        Koman A., Chassin D., and Bellet D.
TITLE          Protein called epil/placentin, process for the preparation of this
                protein and pharmaceutical composition containing such, DNA coding
                for said Protein
JOURNAL        Patent: US 5910480-A 3 08-JUN-1999;
FEATURES       Location/Qualifiers
Source         1..615
BASE COUNT     186 a 145 c 142 g 142 t
ORIGIN
Query Match    100.0%; Score 342; DB 6; Length 615;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAAACACTTGTGTCATATATGCCCCATG 60
Db 182 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAAACACTTGTGTCATATATGCCCCATG 241
Qy 61 CCTGAGNAGACATTCACCCACCACCCAGAGGGTGGCTGCTGGAATCTGGAGTCCCAA 120
Db 242 CCTGAGNAGACATTCACCCACCACCCAGAGGGTGGCTGCTGGAATCTGGAGTCCCAA 301
Qy 121 GAAATGGTGTCAACCTCCCAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180
Db 302 GAAATGGTGTCAACCTCCCAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 361
Qy 181 ATTCCTAATTTGTACAGAGCTGGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 240
Db 362 ATTCCTAATTTGTACAGAGCTGGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 421
Qy 241 AAAATAATATCTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATCTGTTGT 300
Db 422 AAAATAATATCTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATCTGTTGT 481
Qy 301 GAAATATATCTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATCTGTTGT 342
Db 482 GAAATATATCTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATCTGTTGT 523

RESULT 4
HUMINSH
LOCUS          HUMINSH          617 bp      mRNA      linear      PRI 20-SEP-1996
DEFINITION     Homo sapiens early placenta insulin-like peptide EPIL (INSL4) mRNA,
                complete cds.
ACCESSION      L34838
VERSION        L34838.1 GI:1220314
KEYWORDS       insulin family; placentin.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 617)  
 Chassin,D., Laurent,A., Janneau,J.L., Berger,R. and Bellet,D.  
 Cloning of a new member of the insulin gene superfamily (INSL4) expressed in human placenta  
 Genomics 29 (2), 465-470 (1995)  
 96115599  
 MEDLINE  
 PUBMED 8666396  
 2 (bases 1 to 617)  
 Komar,Aetal.  
 Patented application, FR 2721033, 13-JUN-1994  
 Unpublished (1994)  
 JOURNAL  
 FEATURES  
 source  
 1. .617  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /cell\_type="trophoblast"  
 /tissue\_type="placenta"  
 /dev\_stage="6 weeks"  
 /tissue\_lib="subtracted cDNA"  
 1. .617  
 /gene="INSL4"  
 106. .525  
 /note="INSL4"  
 /note="early placenta insulin-like peptide"  
 /codon\_start=1  
 /product="Epil"  
 /protein\_id="AAB08516.1"  
 /db\_xref="GI:1220315"  
 /translation="MASLFRSYLPALWLLSOLLRESLAELRGCGPRFGKLLSYCP  
 MPKFTFTTGGWLLSGRPMKEMVSTNNKDGALGTSFIPNLSPKLPSEGP  
 SLKKIILSRKSRGRHDFPCCCEVICDDGTSVKLCT"  
 BASE COUNT 189 a 144 c 142 g 142 t  
 ORIGIN  
 Query Match 100.0%; Score 342; DB 9; Length 617;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-93;  
 Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACATTCGTGCATATTCGCCCATG 60  
 Db 181 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACATTCGTGCATATTCGCCCATG 240  
 Qy 61 CCTGAGAAGACATTCACACACCCAGGAGGTGGCTGGAACTCTGGACGTCCCAAA 120  
 Db 241 CCTGAGAAGACATTCACACACCCAGGAGGTGGCTGGAACTCTGGACGTCCCAAA 300  
 Qy 121 GAATGTGTCAACCTCCAAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180  
 Db 301 GAATGTGTCAACCTCCAAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 360  
 Qy 181 ATTCCTAATTTGTACAGAGCTGAAGAAACCACTGCTGAAGGGCAGCCATCATTTGAAG 240  
 Db 361 ATTCCTAATTTGTACAGAGCTGAAGAAACCACTGCTGAAGGGCAGCCATCATTTGAAG 420  
 Qy 241 AAAATAATACCTTTCCCGAAAAAGAGAGTGGAGCTCACAGATTTGATCCATTCGTGTG 300  
 Db 421 AAAATAATACCTTTCCCGAAAAAGAGAGTGGAGCTCACAGATTTGATCCATTCGTGTG 480  
 Qy 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 342  
 Db 481 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 522  
 RESULT 5  
 A47856  
 LOCUS  
 DEFINITION Sequence 1 from Patent WO9534653.  
 A47856  
 ACCESSION  
 VERSION A47856.1 GI:2301742  
 KEYWORDS

unidentified  
 unidentified  
 unclassified.  
 1 (bases 1 to 618)  
 Komar,A., Chassin,D. and Bellet,D.  
 EPIL/PLACENTINE  
 Patent: WO 9534653-A 1 21-DEC-1995;  
 ROUSSY INST GUSTAVE (FR)  
 Other publication FR 2721033 951215.  
 COMMENT  
 FEATURES  
 source  
 1. .618  
 Location/Qualifiers  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"  
 107. .526  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAA03084.1"  
 /db\_xref="GI:2301743"  
 /translation="MASLFRSYLPALWLLSOLLRESLAELRGCGPRFGKLLSYCP  
 MPEKFTFTTGGWLLSGRPMKEMVSTNNKDGALGTSFIPNLSPKLPSEGP  
 SLKKIILSRKSRGRHDFPCCCEVICDDGTSVKLCT"  
 BASE COUNT 189 a 145 c 142 g 142 t  
 ORIGIN  
 Query Match 100.0%; Score 342; DB 6; Length 618;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-93;  
 Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACATTCGTGCATATTCGCCCATG 60  
 Db 182 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACATTCGTGCATATTCGCCCATG 241  
 Qy 61 CCTGAGAAGACATTCACACACCCAGGAGGTGGCTGGAACTCTGGACGTCCCAAA 120  
 Db 242 CCTGAGAAGACATTCACACACCCAGGAGGTGGCTGGAACTCTGGACGTCCCAAA 301  
 Qy 121 GAATGTGTCAACCTCCAAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180  
 Db 302 GAATGTGTCAACCTCCAAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 361  
 Qy 181 ATTCCTAATTTGTACAGAGCTGAAGAAACCACTGCTGAAGGGCAGCCATCATTTGAAG 240  
 Db 362 ATTCCTAATTTGTACAGAGCTGAAGAAACCACTGCTGAAGGGCAGCCATCATTTGAAG 421  
 Qy 241 AAAATAATACCTTTCCCGAAAAAGAGAGTGGAGCTCACAGATTTGATCCATTCGTGTG 300  
 Db 422 AAAATAATACCTTTCCCGAAAAAGAGAGTGGAGCTCACAGATTTGATCCATTCGTGTG 481  
 Qy 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 342  
 Db 482 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 523  
 RESULT 6  
 A98978  
 LOCUS  
 DEFINITION Sequence 1 from Patent WO9909172.  
 A98978  
 ACCESSION  
 VERSION A98978.1 GI:6781938  
 KEYWORDS  
 source  
 unidentified  
 unidentified  
 unclassified.  
 1 (bases 1 to 618)  
 Mock,P. and Bellet,D.  
 METHOD FOR IDENTIFYING AND LOCATING EXPRESSED EPIL PEPTIDES, CODED  
 BY THE INSL4 GENE AND THEIR USES  
 Patent: WO 9909172-A 1 25-FEB-1999;  
 MOCK PASCAL (CH); ROUSSY INST GUSTAVE (FR)  
 JOURNAL  
 FEATURES  
 source  
 1. .618  
 Location/Qualifiers  
 /organism="unidentified"

/mol\_type="genomic DNA"  
/db\_xref="taxon:32644"

## CDS

107..526  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CA70586.1"

/db\_xref="GI:6781939"  
/translation="MASLFRSLPAIWLILLSOLLRESLAELRGCGPRFGKHLISYCP  
MPKFTFTTGGWLLSGRPEKEMVSTNNKDGALGTTSEFIPNLSPCLKPLSEQP  
SLKIIILSRKSRGHRFPFCCEVICDDGTSVKLCT"

BASE COUNT 189 a 145 c 142 g 142 t

## ORIGIN

Query Match 100.0%; Score 342; DB 6; Length 618;

Best Local Similarity 100.0%; Pred. No. 1.3e-93;

Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGGTCCCGCATTTGGAACACACTTCTGTCATATTGCCCATG 60

DB 182 GCAGAGCTGAGGGGATGGTCCCGCATTTGGAACACACTTCTGTCATATTGCCCATG 241

QY 61 CCTGAGAAGACATTCACCCACCACCCAGGAGGTGGCTGCTGAATCTGGAGCTGCCAAA 120

DB 242 CCTGAGAAGACATTCACCCACCACCCAGGAGGTGGCTGCTGAATCTGGAGCTGCCAAA 301

QY 121 GAAATGGTGTCAACCTCCAAACAACAAGATGGACAGCCCTTAGGTACGACATCAGAATTC 180

DB 302 GAAATGGTGTCAACCTCCAAACAACAAGATGGACAGCCCTTAGGTACGACATCAGAATTC 361

QY 181 ATTCTTAATTTGTCCAGAGCTGAAGAACCACTCTCTGAAGGGCAGGCATCATTTGAAG 240

DB 362 ATTCTTAATTTGTCCAGAGCTGAAGAACCACTCTCTGAAGGGCAGGCATCATTTGAAG 421

QY 241 AAATAATATCTTTCCCGCAAAAAGAGAGTGGACGTCACAGATTTGATCCATCTCTTGT 300

DB 422 AAATAATATCTTTCCCGCAAAAAGAGAGTGGACGTCACAGATTTGATCCATCTCTTGT 481

QY 301 GAAGTAATTTGTGACGATGGAATTCAGTTAAATTATGTACA 342

DB 482 GAAGTAATTTGTGACGATGGAATTCAGTTAAATTATGTACA 523

RESULT 7  
AX019983  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Laurent, A. and Bellet, D.  
Expression of the insl4 gene in human embryonic bone tissues and  
applications  
PATENT: WO 9337780-A 1 29-JUL-1999;  
ROUSSY INST GUSTAVE (FR); LAURENT ANNE (FR); BELLET DOMINIQUE (FR)  
FEATURES  
source  
1..618  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
107..526  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC07742.1"  
/db\_xref="GI:10043813"  
/translation="MASLFRSLPAIWLILLSOLLRESLAELRGCGPRFGKHLISYCP  
MPKFTFTTGGWLLSGRPEKEMVSTNNKDGALGTTSEFIPNLSPCLKPLSEQP  
SLKIIILSRKSRGHRFPFCCEVICDDGTSVKLCT"

BASE COUNT 189 a 145 c 142 g 142 t

## ORIGIN

Query Match 100.0%; Score 342; DB 6; Length 618;

Best Local Similarity 100.0%; Pred. No. 1.3e-93;

Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGGTCCCGCATTTGGAACACACTTCTGTCATATTGCCCATG 60

DB 182 GCAGAGCTGAGGGGATGGTCCCGCATTTGGAACACACTTCTGTCATATTGCCCATG 241

QY 61 CCTGAGAAGACATTCACCCACCACCCAGGAGGTGGCTGCTGAATCTGGAGCTGCCAAA 120

DB 242 CCTGAGAAGACATTCACCCACCACCCAGGAGGTGGCTGCTGAATCTGGAGCTGCCAAA 301

QY 121 GAAATGGTGTCAACCTCCAAACAACAAGATGGACAGCCCTTAGGTACGACATCAGAATTC 180

DB 302 GAAATGGTGTCAACCTCCAAACAACAAGATGGACAGCCCTTAGGTACGACATCAGAATTC 361

QY 181 ATTCTTAATTTGTCCAGAGCTGAAGAACCACTCTCTGAAGGGCAGGCATCATTTGAAG 240

DB 362 ATTCTTAATTTGTCCAGAGCTGAAGAACCACTCTCTGAAGGGCAGGCATCATTTGAAG 421

QY 241 AAATAATATCTTTCCCGCAAAAAGAGAGTGGACGTCACAGATTTGATCCATCTCTTGT 300

DB 422 AAATAATATCTTTCCCGCAAAAAGAGAGTGGACGTCACAGATTTGATCCATCTCTTGT 481

QY 301 GAAGTAATTTGTGACGATGGAATTCAGTTAAATTATGTACA 342

DB 482 GAAGTAATTTGTGACGATGGAATTCAGTTAAATTATGTACA 523

BD107918 618 bp DNA linear PAT 18-SEP-2002  
Expression of the INSL4 gene in human embryonic bone tissues and  
applications.  
BD107918  
BD107918.1 GI:23202736  
JP 2002500888-A/1.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 618)  
Laurent, A. and Bourez, D.  
Expression of the INSL4 gene in human embryonic bone tissues and  
applications  
Patent: JP 2002500888-A 1 15-JAN-2002;  
INSTITUT GUSTAVE ROUSSY  
OS Homo sapiens (human)  
PN JP 2002500888-A/1  
PD 15-JAN-2002  
PP 22-JAN-1999 JP 2000528687  
PR 23-JAN-1998 FR 98/00715  
PI ANN LAURENT, DOMINIQUE BOUREZ  
PC C12N15/09, A61K39/395, A61K39/00, A61K45/00, A61K48/00,  
PC A61P19/00,  
PC C07H21/00, C07K14/64, C07K16/26, C12N5/10, C12P21/02, C12Q1/68, PC  
GOIN33/53  
PC GOIN33/569, GOIN33/68, C12N15/00, A61K37/02, C12N5/00 CC  
Expression of the INSL4 gene in human embryonic bone tissues CC  
and  
CC applications Location/Qualifiers  
FH Key (107)..(523).  
FT CDS Location/Qualifiers  
1..618  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
189 a 145 c 142 g 142 t

FEATURES  
source

BASE COUNT 189 a 145 c 142 g 142 t  
ORIGIN

```

Query Match      100.0%; Score 342; DB 6; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAACACTTGTGTCATATTCGCCCATG 60
   |||||||
DB 182 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAACACTTGTGTCATATTCGCCCATG 241
   |||||||

QY 61 CTTGAGAAGACATTCACCCACCCAGGAGGTTGCTGCTGAATCTGGACGTCCTCCAAA 120
   |||||||
DB 242 CTTGAGAAGACATTCACCCACCCAGGAGGTTGCTGCTGAATCTGGACGTCCTCCAAA 301
   |||||||

QY 121 GAAATGGTGTCAACCTCCAAACAAGATGGAGACCTTAGGTACGACATCAGAATTC 180
   |||||||
DB 302 GAAATGGTGTCAACCTCCAAACAAGATGGAGACCTTAGGTACGACATCAGAATTC 361
   |||||||

QY 181 ATTCTTAATTTGACAGAGCTGAAGAAACCACTGCTGCTGAAGGGCAGCCATCATTTGAAG 240
   |||||||
DB 362 ATTCTTAATTTGACAGAGCTGAAGAAACCACTGCTGCTGAAGGGCAGCCATCATTTGAAG 421
   |||||||

QY 241 AAAATAATACCTTTCCCGCAAAAAGAGAGTGGAGCTCACAGATTTGATCAATTCCTTGT 300
   |||||||
DB 422 AAAATAATACCTTTCCCGCAAAAAGAGAGTGGAGCTCACAGATTTGATCAATTCCTTGT 481
   |||||||

QY 301 GAAGTAATTTGACGATGGAATTCAGTTAAATTAATGTACA 342
   |||||||
DB 482 GAAGTAATTTGACGATGGAATTCAGTTAAATTAATGTACA 523
   |||||||

RESULT 9
BC026254
LOCUS      Homo sapiens, insulin-like 4 (placenta), clone MGC:22499
DEFINITION IMAGE:4770395, mRNA, complete cds.
ACCESSION BC026254
VERSION    BC026254.1 GI:20070772
KEYWORDS   MGC.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Strausberg, R.
Direct Submission
Submitted (02-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAL Plate: 37 Row: m Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504710.
Location/Qualifiers
1. .673
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="locusID:3641"
/db_xref="taxon:9606"
/clone="MGC:22499 IMAGE:4770395"

```

```

/tissue_type="Placenta"
/clone_lib="NIH_MGC_79"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"
123. .542
/codon_start=1
/product="insulin-like 4 (placenta)"
/protein_id="AAH26254.1"
/db_xref="GI:20070772"
/translation="MASLFRSYLPAIMLLSOLLRESLAARLRCGPRFKHLISYCP
MPKFTFTTGGWLLLESGRPKEMVSTNNKGOALGTTSEFIPNLSPELKPLSEGP
SLKILLSRKKRSRHRDPFCFEVCDGTSYKLCI"
BASE COUNT 228 a 149 c 150 g 146 t
ORIGIN

Query Match      100.0%; Score 342; DB 9; Length 673;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAACACTTGTGTCATATTCGCCCATG 60
   |||||||
DB 198 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAACACTTGTGTCATATTCGCCCATG 257
   |||||||

QY 61 CTTGAGAAGACATTCACCCACCCAGGAGGTTGCTGCTGAATCTGGAGCTGCCAAA 120
   |||||||
DB 258 CTTGAGAAGACATTCACCCACCCAGGAGGTTGCTGCTGAATCTGGAGCTGCCAAA 317
   |||||||

QY 121 GAAATGGTGTCAACCTCCAAACAAGATGGAGACCTTAGGTACGACATCAGAATTC 180
   |||||||
DB 318 GAAATGGTGTCAACCTCCAAACAAGATGGAGACCTTAGGTACGACATCAGAATTC 377
   |||||||

QY 181 ATTCTTAATTTGACAGAGCTGAAGAAACCACTGCTGCTGAAGGGCAGCCATCATTTGAAG 240
   |||||||
DB 378 ATTCTTAATTTGACAGAGCTGAAGAAACCACTGCTGCTGAAGGGCAGCCATCATTTGAAG 437
   |||||||

QY 241 AAAATAATACCTTTCCCGCAAAAAGAGAGTGGAGCTCACAGATTTGATCCATTCCTTGT 300
   |||||||
DB 438 AAAATAATACCTTTCCCGCAAAAAGAGAGTGGAGCTCACAGATTTGATCCATTCCTTGT 497
   |||||||

QY 301 GAAGTAATTTGACGATGGAATTCAGTTAAATTAATGTACA 342
   |||||||
DB 498 GAAGTAATTTGACGATGGAATTCAGTTAAATTAATGTACA 539
   |||||||

RESULT 10
AR126781
LOCUS      Sequence 4 from patent US 6180364.
DEFINITION AR126781
ACCESSION  AR126781
VERSION    AR126781.1 GI:14113374
KEYWORDS   .
SOURCE     Unknown.
ORGANISM   Unclassified.
REFERENCE  1 (bases 1 to 934)
AUTHORS   Koman,A., Chassin,D. and Bellet,D.
TITLE     Protein called epil/placenta, process for the preparation of this
          protein and pharmaceutical composition containing such, DNA coding
          for said protein
JOURNAL   Patent: US 6180364-A 4 30-JAN-2001;
FEATURES   Location/Qualifiers
           source
           1. .934
           /organism="unknown"
BASE COUNT 311 a 188 c 226 g 209 t
ORIGIN

Query Match      100.0%; Score 342; DB 6; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.4e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAACACTTGTGTCATATTCGCCCATG 60
   |||||||
DB 501 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAACACTTGTGTCATATTCGCCCATG 560
   |||||||

```

QY 61 CTGAGAGACATTACACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTCCCAAA 120  
|||||  
Db 561 CTTGAGAGACATTACACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTCCCAAA 620  
|||||  
QY 121 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCTTAGGTACGACATCAGAATTC 180  
|||||  
Db 621 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCTTAGGTACGACATCAGAATTC 680  
|||||  
QY 181 ATTCTTAATTTGTACACAGAGCTGAAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 240  
|||||  
Db 681 ATTCTTAATTTGTACACAGAGCTGAAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 740  
|||||  
QY 241 AAAATAATTAATTTCCCGCAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 300  
|||||  
Db 741 AAAATAATTAATTTCCCGCAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 800  
|||||  
QY 301 GAAATTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA 342  
|||||  
Db 801 GAAATTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA 842  
|||||

RESULT 11  
AR202438  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

Query Match 100.0%; Score 342; DB 6; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.4e-93;  
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTCAATATGCCCCATG 60  
|||||  
Db 501 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTCAATATGCCCCATG 560  
|||||  
QY 61 CTGAGAGACATTACACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTCCCAAA 120  
|||||  
Db 561 CTTGAGAGACATTACACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTCCCAAA 620  
|||||  
QY 121 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCTTAGGTACGACATCAGAATTC 180  
|||||  
Db 621 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCTTAGGTACGACATCAGAATTC 680  
|||||  
QY 181 ATTCTTAATTTGTACACAGAGCTGAAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 240  
|||||  
Db 681 ATTCTTAATTTGTACACAGAGCTGAAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 740  
|||||  
QY 241 AAAATAATTAATTTCCCGCAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 300  
|||||  
Db 741 AAAATAATTAATTTCCCGCAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 800  
|||||  
QY 301 GAAATTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA 342  
|||||  
Db 801 GAAATTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA 842  
|||||

Query Match 100.0%; Score 342; DB 6; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.4e-93;  
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTCAATATGCCCCATG 60  
|||||  
Db 501 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTCAATATGCCCCATG 560  
|||||  
QY 61 CTGAGAGACATTACACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTCCCAAA 120  
|||||  
Db 561 CTTGAGAGACATTACACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTCCCAAA 620  
|||||  
QY 121 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCTTAGGTACGACATCAGAATTC 180  
|||||  
Db 621 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCTTAGGTACGACATCAGAATTC 680  
|||||  
QY 181 ATTCTTAATTTGTACACAGAGCTGAAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 240  
|||||  
Db 681 ATTCTTAATTTGTACACAGAGCTGAAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 740  
|||||  
QY 241 AAAATAATTAATTTCCCGCAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 300  
|||||  
Db 741 AAAATAATTAATTTCCCGCAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 800  
|||||  
QY 301 GAAATTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA 342  
|||||  
Db 801 GAAATTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA 842  
|||||

RESULT 12  
AR071299  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

Query Match 99.5%; Score 340.4; DB 6; Length 615;  
Best Local Similarity 99.7%; Pred. No. 4.2e-93;  
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTCAATATGCCCCATG 60  
|||||  
Db 181 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTCAATATGCCCCATG 240  
|||||  
QY 61 CTGAGAGACATTACACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTCCCAAA 120  
|||||  
Db 241 CTGAGAGACATTACACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTCCCAAA 300  
|||||  
QY 121 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCTTAGGTACGACATCAGAATTC 180  
|||||  
Db 301 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCTTAGGTACGACATCAGAATTC 360  
|||||  
QY 181 ATTCTTAATTTGTACACAGAGCTGAAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 240  
|||||  
Db 361 ATTCTTAATTTGTACACAGAGCTGAAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 420  
|||||  
QY 241 AAAATAATTAATTTCCCGCAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 300  
|||||  
Db 421 AAAATAATTAATTTCCCGCAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 480  
|||||  
QY 301 GAAATTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA 342  
|||||  
Db 481 GAAATTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA 522  
|||||

RESULT 13  
AR126780  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

Query Match 100.0%; Score 342; DB 6; Length 615;  
Best Local Similarity 100.0%; Pred. No. 1.4e-93;  
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTCAATATGCCCCATG 60  
|||||  
Db 501 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTCAATATGCCCCATG 560  
|||||  
QY 61 CTGAGAGACATTACACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTCCCAAA 120  
|||||  
Db 561 CTTGAGAGACATTACACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTCCCAAA 620  
|||||  
QY 121 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCTTAGGTACGACATCAGAATTC 180  
|||||  
Db 621 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCTTAGGTACGACATCAGAATTC 680  
|||||  
QY 181 ATTCTTAATTTGTACACAGAGCTGAAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 240  
|||||  
Db 681 ATTCTTAATTTGTACACAGAGCTGAAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 740  
|||||  
QY 241 AAAATAATTAATTTCCCGCAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 300  
|||||  
Db 741 AAAATAATTAATTTCCCGCAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 800  
|||||  
QY 301 GAAATTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA 342  
|||||  
Db 801 GAAATTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA 522  
|||||

Query Match 99.5%; Score 340.4; DB 6; Length 615;  
Best Local Similarity 99.7%; Pred. No. 4.2e-93;  
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATTGCCCATG 60  
|||||  
DB 181 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATTGCCCATG 240

QY 61 CTTGAGAGACATTCACCCACCACCCAGGAGGTGGCTGCGAATCTGGAGCTCCCAA 120  
|||||  
DB 241 CTTGAGAGACATTCACCCACCACCCAGGAGGTGGCTGCGAATCTGGAGCTCCCAA 300

QY 121 GAAATGGTGTCAACCTCCAAACAAGATGGCAAGCCCTTAGGTAGCAGCATCAGAATTC 180  
|||||  
DB 301 GAAATGGTGTCAACCTCCAAACAAGATGGCAAGCCCTTAGGTAGCAGCATCAGAATTC 360

QY 181 ATTCTTAATTTGTACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 240  
|||||  
DB 361 ATTCTTAATTTGTACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 420

QY 241 AAAATAATACCTTCCCGCAAAAGAGAGTGGACGTACAGATTTGATCCATTCCTGTGT 300  
|||||  
DB 421 AAAATAATACCTTCCCGCAAAAGAGAGTGGACGTACAGATTTGATCCATTCCTGTGT 480

QY 301 GAAGTAATTTGTGACGATGGAATTCAGTTAAATTTATGTACA 342  
|||||  
DB 481 GAAGTAATTTGTGACGATGGAATTCAGTTAAATTTATGTACA 522

RESULT 14  
AR202437 AR202437 615 bp DNA linear PAT 20-APR-2002  
LOCUS Sequence 1 from patent US 6362318.  
DEFINITION AR202437  
ACCESSION AR202437  
VERSION AR202437.1 GI:20256976  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 615)  
AUTHORS Koman,A., Chassin,D. and Bellet,D.  
TITLE Protein called EpiT/placentin, process for the preparation of this protein and pharmaceutical composition containing such, DNA coding for said protein  
JOURNAL Patent: US 6362318-A 1 26-MAR-2002;  
FEATURES Location/Qualifiers  
source 1..615  
BASE COUNT 187 a 143 c 143 g 142 t

Query Match 99.5%; Score 340.4; DB 6; Length 615;  
Best Local Similarity 99.7%; Pred. No. 4.2e-93;  
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATTGCCCATG 60  
|||||  
DB 181 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATTGCCCATG 240

QY 61 CTTGAGAGACATTCACCCACCACCCAGGAGGTGGCTGCGAATCTGGAGCTCCCAA 120  
|||||  
DB 241 CTTGAGAGACATTCACCCACCACCCAGGAGGTGGCTGCGAATCTGGAGCTCCCAA 300

QY 121 GAAATGGTGTCAACCTCCAAACAAGATGGCAAGCCCTTAGGTAGCAGCATCAGAATTC 180  
|||||  
DB 301 GAAATGGTGTCAACCTCCAAACAAGATGGCAAGCCCTTAGGTAGCAGCATCAGAATTC 360

QY 181 ATTCTTAATTTGTACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 240  
|||||  
DB 361 ATTCTTAATTTGTACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 420

QY 241 AAAATAATACCTTCCCGCAAAAGAGAGTGGACGTACAGATTTGATCCATTCCTGTGT 300  
|||||

Db 421 AAAATAATACCTTCCCGCAAAAGAGAGTGGACGTACAGATTTGATCCATTCCTGTGT 480

QY 301 GAAGTAATTTGTGACGATGGAATTCAGTTAAATTTATGTACA 342  
|||||  
DB 481 GAAGTAATTTGTGACGATGGAATTCAGTTAAATTTATGTACA 522

RESULT 15  
G27877/c G27877 482 bp DNA linear STS 29-JUN-1996  
LOCUS human STS SHGC-34067, sequence tagged site.  
DEFINITION G27877  
ACCESSION G27877  
VERSION G27877.1 GI:1396596  
KEYWORDS STS; STS sequence; primer; sequence tagged site.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 482)  
AUTHORS Myers,R.M.  
JOURNAL Unpublished (1996)  
COMMENT  
Contact: Richard M. Myers  
Stanford Human Genome Center (SHGC)  
Stanford University School of Medicine  
Department of Genetics, M-344, Stanford, CA 94305, USA  
Tel: 4157259687  
Fax: 4157259689  
Email: myers@shgc.stanford.edu

Primer A: TCAGTGAATTTGTCATTGAGAATAC  
Primer B: AAAGAGAAGTGGAGTCCACAG  
STS size: 150  
PCR Profile:  
Initial incubation: 94 degrees C for 90 seconds  
Denaturation: 94 degrees C for 15 seconds  
Annealing: 62 degrees C for 23 seconds  
Polymerization: 72 degrees C for 30 seconds  
PCR Cycles: 30  
Thermal Cycler: Perkin Elmer 9600  
Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
Tag Polymerase: 0.05 units/ul  
Total Vol: 10 ul

Buffer:  
MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 20 mM  
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from R75651  
-- Washington University/Merck EST sequence.

FEATURES  
source 1..482  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="9"  
STS  
primer\_bind 1..150  
primer\_bind complement(130..150)  
BASE COUNT 114 a 99 c 121 g 146 t 2 others  
ORIGIN

Query Match 70.5%; Score 241.2; DB 11; Length 482;  
Best Local Similarity 91.5%; Pred. NO. 1e-62;  
Matches 324; Conservative 0; Mismatches 18; Indels 12; Gaps 6;

QY 1 GCAGAGCTGAGGGGATGTGTCGCC---GATTGGAAACACTTGTCTGTC---ATATTGC 54

422	GCAAACTCGAGGGATGTGTCGCCCGAATTTGGGAAACACTGTGTCCTATATTGCC	363
55	CCATGCTCGAGAAGACATT---CACCACCACCCAGGAGGGTGGCTGCTGGAATCTGGA	111
362	CCATGCCCTGAGAAGACATTCACCCACCACCCAGGAGGGTGGCTGCTGGAATCTGGA	303
112	CGT-CCCAAAGAAATGGTGTCAACCT-CCAAACAACAAGATGACAAG-CTTTAGGTACG	168
302	CGTCCCAAAGAAATGGTGTCAACCTCCCAACAACAAGATGACAAGCCCTTAGGTACG	243
169	ACATCAGAAATCATTCCTAAATTTGTCACCAGAGCTGAGAAACCACTGTCGAAGGGCAG	228
242	ACATCAGAAATCATTCCTAAATTTGTCACCAGAGCTGAAGAAACCACTGTCGAAGGGCAG	183
229	CCATCATTTGAAGAAAATAATACTTTCCCGCAAAAGAGAGTGGAGCTCACAGATTTGAT	288
182	CCATCATTTGAAGAAAATAATACTTTCCCGCAAAAGAGAGTGGAGCTCACAGATTTGAT	123
289	CCATCTGTGTGGAAGTAATTTGTGACGATGGAACTTCAGTTAAATNTGTACA	342
122	CCATTCTGTGTGGAAGTAATTTGTGACGATGGAACTTCAGTTAAATNTGTACA	69

Search completed: October 17, 2003, 02:21:35  
Job time : 1737 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2003, 23:31:46 ; Search time 206 seconds

(without alignments)  
4481.591 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 342

Sequence: 1 GCAGAGCTGAGGGCATGTGG.....CTTCAGTTAAATTATGTACA 342

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_19Jun03:\*  
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	342	100.0	420	19 AAV44663	Zinsl coding seque
2	342	100.0	480	19 AAV44664	N-terminally tagge
3	342	100.0	618	17 AAT10275	Human early placen
4	342	100.0	618	20 AAX83561	Human insulin-like
5	342	100.0	618	20 AAX27490	Human insulin-like
6	342	100.0	649	18 AAT68419	Human relaxin-rela
7	84.8	24.8	752	4 AAN30186	Sequence encoding
8	84.8	24.8	752	4 AAN30196	Sequence encoding

C	9	81.4	23.8	2364	23 AAS69640	DNA encoding novel
C	10	81.4	23.8	2364	23 AAS73158	DNA encoding novel
	11	71.8	21.0	475	23 ABV42732	Human prostate exp
	12	71.8	21.0	489	11 AAQ06595	H2 prorelaxin gene
	13	71.8	21.0	495	11 AAQ06596	H2 prorelaxin gene
	14	71.8	21.0	558	5 AAN40086	Sequence encoding
	15	71.8	21.0	558	10 AAN92483	mRNA sequence of h
	16	71.8	21.0	915	16 AAQ76309	Not I-Bam HI fragm
	17	71.8	21.0	1120	23 ABV22241	Human prostate exp
	18	71.8	21.0	1120	23 ABV22244	Human prostate exp
	19	71.8	21.0	1120	23 ABV22256	Human prostate exp
	20	71.8	21.0	1120	23 ABV22268	Human prostate exp
	21	71.8	21.0	1120	23 ABV22272	Human prostate exp
	22	71.8	21.0	1120	23 ABV22275	Human prostate exp
	23	71.8	21.0	1120	23 ABV22277	Human prostate exp
	24	71.8	21.0	1120	23 ABV22278	Human prostate exp
	25	71.8	21.0	1120	23 ABV22280	Human prostate exp
	26	71.8	21.0	1120	23 ABV22281	Human prostate exp
	27	71.8	21.0	1120	23 ABV22284	Human prostate exp
	28	71.8	21.0	1120	23 ABV22285	Human prostate exp
	29	71.8	21.0	1120	23 ABV22286	Human prostate exp
	30	71.8	21.0	1120	23 ABV22287	Human prostate exp
	31	71.8	21.0	1120	23 ABV22288	Human prostate exp
	32	71.8	21.0	1120	23 ABV22289	Human prostate exp
	33	71.8	21.0	1120	23 ABV22290	Human prostate exp
	34	71.8	21.0	1120	23 ABV22292	Human prostate exp
	35	71.8	21.0	1120	23 ABV22298	Human prostate exp
	36	71.8	21.0	1120	23 ABV22304	Human prostate exp
	37	71.8	21.0	1120	23 ABV22305	Human prostate exp
	38	71.8	21.0	1120	23 ABV22315	Human prostate exp
	39	71.8	21.0	1120	23 ABV22317	Human prostate exp
	40	71.8	21.0	1120	23 ABV22321	Human prostate exp
	41	71.8	21.0	1120	23 ABV22323	Human prostate exp
	42	71.8	21.0	1120	23 ABV22332	Human prostate exp
	43	71.8	21.0	1120	23 ABV22333	Human prostate exp
	44	71.8	21.0	1120	23 ABV22335	Human prostate exp
	45	71.8	21.0	1120	23 ABV22336	Human prostate exp

#### ALIGNMENTS

RESULT 1  
AAV44663  
ID AAV44663 standard; cDNA; 420 bp.  
XX  
AC AAV44663;  
XX  
DT 07-OCT-1998 (first entry)  
XX  
DE Zinsl coding sequence.  
XX  
KW Zinsl; human; placenta; pancreatic islet cell proliferation;  
KW Insulin secretion; diabetes; therapy; ds.  
XX  
OS Homo sapiens.  
XX  
PN W09827210-A1.  
XX  
PD 25-JUN-1998.  
XX  
PF 16-DEC-1997; 97WO-US33326.  
XX  
PR 16-DEC-1996; 96US-0033003.  
XX  
PR (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Conklin DC, Hoffmann RC, Humes JM, Jaspers SR, Ren HP;  
PI Sprugel KH;  
XX  
DR WPI; 1998-362779/31.  
DR P-PSDB; AAW69168.  
XX

PT New isolated protein, Zinsl - obtained from human placenta, which  
PT increases the proliferation of pancreatic islet cells, used for  
XX treating diabetes  
XX  
XX Claim 1; Page 60-61; 77pp; English.  
XX  
CC This sequence encodes the human Zinsl protein of the invention. The Zinsl  
CC protein was isolated from human placenta, and is believed to be a new  
CC version of the mature protein of placentalin, having disulphide bonded A  
CC and B chains. The protein can be used for stimulating the proliferation  
CC of pancreatic islets to increase insulin secretory capacity of mammals.  
CC In particular it can be used for the treatment of diabetes. It can also  
CC be used for stimulating in vitro proliferation of pancreatic islet cells.  
CC It can also be used for production of antibodies and in detection and  
CC diagnosis.  
XX  
SQ Sequence 420 BP; 124 A; 100 C; 98 G; 98 T; 0 other;  
Query Match 100.0%; Score 342; DB 19; Length 420;  
Best Local Similarity 100.0%; Pred. No. 9.2e-101;  
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCAGAGCTGAGGGGATGGTCCCGATTGGAAACACTTGTCTCATATTGCCCATG 60  
DB 76 GCAGAGCTGAGGGGATGGTCCCGATTGGAAACACTTGTCTCATATTGCCCATG 135  
QY 61 CCTGAGAGACATTCACACACCCAGGAGGGTGGCTGGATCTGGAGCTGCCAAA 120  
DB 136 CCTGAGAGACATTCACACACCCAGGAGGGTGGCTGGATCTGGAGCTGCCAAA 195  
QY 121 GAAATGGTCTCAACCTCCCAACCAAGAGTGGACAGCCCTTAGGTACGACATCAGAAATTC 180  
DB 196 GAAATGGTCTCAACCTCCCAACCAAGAGTGGACAGCCCTTAGGTACGACATCAGAAATTC 255  
QY 181 ATTCCTAATTTGTACACAGCTGAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 240  
DB 256 ATTCCTAATTTGTACACAGCTGAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 315  
QY 241 AAAATAAATCTTTCCCGCAAAAGAGAGTGGACGCTACAGATTTGATCCCATTCGTGTGT 300  
DB 316 AAAATAAATCTTTCCCGCAAAAGAGAGTGGACGCTACAGATTTGATCCCATTCGTGTGT 375  
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 342  
DB 376 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 417  
RESULT 2  
AAV44664  
ID AAV44664 standard; cDNA; 480 BP.  
XX  
XX AAV44664;  
AC  
XX  
DT 07-OCT-1998 (first entry)  
XX  
DE N-terminally tagged Zinsl coding sequence.  
XX  
KW Zinsl; human; placentalin; pancreatic islet cell proliferation;  
KW insulin secretion; diabetes; therapy; ds.  
XX  
OS Homo sapiens.  
XX  
PN W09827210-A1.  
XX  
PD 25-JUN-1998.  
XX  
PF 16-DEC-1997; 97WO-US23326.  
XX  
PR 16-DEC-1996; 96US-0033003.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Conklin DC, Hoffmann RC, Humes JM, Jaspers SR, Ren HP;

PI Sprugel KH;  
XX  
DR WPI; 1998-362779/31.  
DR P-PSDB; AAW69170.  
XX  
PT New isolated protein, Zinsl - obtained from human placenta, which  
PT increases the proliferation of pancreatic islet cells, used for  
XX treating diabetes  
XX  
XX Example 1; Page 62-63; 77pp; English.  
XX  
CC This sequence encodes a N-terminally tagged version of the human Zinsl  
CC protein of the invention. The Zinsl protein was isolated from human  
CC placenta, and is believed to be a new version of the mature protein of  
CC placentalin, having disulphide bonded A and B chains. The protein can be  
CC used for stimulating the proliferation of pancreatic islets to increase  
CC insulin secretory capacity of mammals. In particular it can be used for  
CC the treatment of diabetes. It can also be used for stimulating in vitro  
CC proliferation of pancreatic islet cells. It can also be used for  
CC production of antibodies and in detection and diagnosis.  
XX  
SQ Sequence 480 BP; 136 A; 114 C; 121 G; 109 T; 0 other;  
Query Match 100.0%; Score 342; DB 19; Length 480;  
Best Local Similarity 100.0%; Pred. No. 9.8e-101;  
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCAGAGCTGAGGGGATGGTCCCGATTGGAAACACTTGTCTCATATTGCCCATG 60  
DB 136 GCAGAGCTGAGGGGATGGTCCCGATTGGAAACACTTGTCTCATATTGCCCATG 195  
QY 61 CCTGAGAGACATTCACACACCCAGGAGGGTGGCTGGATCTGGAGCTGCCAAA 120  
DB 196 CCTGAGAGACATTCACACACCCAGGAGGGTGGCTGGATCTGGAGCTGCCAAA 255  
QY 121 GAAATGGTCTCAACCTCCCAACCAAGAGTGGACAGCCCTTAGGTACGACATCAGAAATTC 180  
DB 256 GAAATGGTCTCAACCTCCCAACCAAGAGTGGACAGCCCTTAGGTACGACATCAGAAATTC 315  
QY 181 ATTCCTAATTTGTGACGAGCTGAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 240  
DB 316 ATTCCTAATTTGTGACGAGCTGAGAAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 375  
QY 241 AAAATAAATCTTTCCCGCAAAAGAGAGTGGACGCTACAGATTTGATCCCATTCGTGTGT 300  
DB 376 AAAATAAATCTTTCCCGCAAAAGAGAGTGGACGCTACAGATTTGATCCCATTCGTGTGT 435  
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 342  
DB 436 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 477  
RESULT 3  
AAT10275  
ID AAT10275 standard; cDNA; 618 BP.  
XX  
XX AAT10275;  
AC  
XX  
DT 30-JUL-1996 (first entry)  
XX  
DE Human early placental insulin-like protein gene.  
XX  
KW Insulin growth hormone family; early placental insulin like protein;  
KW placentalin; placental tissue; cytotrophoblast; trimer; probe; primer;  
KW amplification; polymerase chain reaction; tyrosine phosphorylation;  
KW cellular protein; growth factor; human; lactation; promoter; PCR;  
KW regeneration; nerve; muscle; skin; bone tissue; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT CDS  
FT Location/Qualifiers  
FT 107..526  
FT /\*tag= a

FT /product= EPIL/placentin  
 FT 107..157  
 FT /\*tag= b  
 FT mat\_peptide 158..523  
 FT /\*tag= c  
 XX  
 PW W09534653-A1.  
 XX  
 XX 21-DEC-1995.  
 XX  
 XX 12-JUN-1995; 95WO-FR00766.  
 XX  
 XX 13-JUN-1994; 94FR-0007191.  
 XX  
 XX (INSR ) INST ROUSSY GUSTAVE.  
 XX  
 XX Bellet D, Chassin D, Koman A;  
 XX  
 XX WPI: 1996-049682/05.  
 DR P-PSDB: AAR89134.  
 XX  
 XX Early placental insulin-like protein, EPIL/placentin - contains  
 PT growth factor-like activity useful for e.g. promoting lactation or  
 PT for regeneration of nerve, muscle, skin or bone tissue  
 XX  
 PS Claim 1; Page 14-15; 25pp; French.

XX This is the nucleotide sequence encoding a novel member of the insulin  
 CC growth hormone family designated early placental insulin like (EPIL)  
 CC protein or placentin. The gene has been found to be expressed  
 CC exclusively in the early placental tissue with a small amount in full  
 CC term placental tissue. No other tissue expresses this gene. The gene  
 CC was obtained from a cDNA library derived from mRNA isolated from  
 CC cytotrophoblasts derived from first trimester placental tissue using,  
 CC as a probe, a fragment of the gene amplified by primers AAT10276-7.  
 CC Although the specific activity of the protein remains to be elucidated,  
 CC it is thought that the protein will induce tyrosine phosphorylation of  
 CC cellular proteins and may have growth factor-like activities e.g. human  
 CC growth factor type 1 or lactation promoter activities. It may also be  
 CC used to regenerate e.g. nerve, muscle, skin or bone tissue.

XX Sequence 618 BP; 189 A; 145 C; 142 G; 142 T; 0 other;

Query Match 100.0%; Score 342; DB 17; Length 618;  
 Best Local Similarity 100.0%; Pred. No. 1.le-100;  
 Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGAGAGCTGAGGGGATGTGTCGCCGGATTTGGAAACACCTTGTCTCATATTGCCCATG 60  
 DB 182 CGAGAGCTGAGGGGATGTGTCGCCGGATTTGGAAACACCTTGTCTCATATTGCCCATG 241  
 QY 61 CCTGAGAAGACATTCACCAACCCAGGAGGGTGGCTGGAACTCTGGACGTCCCAAA 120  
 DB 242 CCTGAGAAGACATTCACCAACCCAGGAGGGTGGCTGGAACTCTGGACGTCCCAAA 301  
 QY 121 GAAATGGTGTCACTCCCAACCAAGATGGACAAGCCTTAGTAGACATCAGAATTC 180  
 DB 302 GAAATGGTGTCACTCCCAACCAAGATGGACAAGCCTTAGTAGACATCAGAATTC 361  
 QY 181 ATTCCTTAATTTGTACAGAGCTGAAGAACCACTGTCTGAAGGCGCCATCATTAAG 240  
 DB 362 ATTCCTTAATTTGTACAGAGCTGAAGAACCACTGTCTGAAGGCGCCATCATTAAG 421  
 QY 241 AAATAAATACTTCCCGCAAAAGAGAGTGGACGTCACAGATTGATTCATTCTGTGT 300  
 DB 422 AAATAAATACTTCCCGCAAAAGAGAGTGGACGTCACAGATTGATTCATTCTGTGT 481  
 QY 301 GAAGTAATTTGTGAGCATGGAACTTCAGTTAAATATGTACA 342  
 DB 482 GAAGTAATTTGTGAGCATGGAACTTCAGTTAAATATGTACA 523

RESULT 4

AA83561  
 ID AAX83561 standard; cDNA: 618 BP.  
 XX  
 AC AAX83561;  
 XX  
 DT 21-DEC-1999 (first entry)  
 XX  
 DE Human insulin-like 4 cDNA sequence.  
 XX  
 KW Human; insulin-like 4; INSL4; embryonic; bone tissue; ligament;  
 KW early-placental insulin-like protein; EPIL; antibody; probe; primer;  
 KW diagnosis; pathology; differentiation; proliferation; cartilage;  
 KW ossification; osteoporosis; dysplasia; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 107..526  
 FT /\*tag= a  
 FT /product= "INSL4"  
 FT /note= "insulin-like 4 protein"  
 XX  
 PN W0937780-A1.  
 XX  
 XX 29-JUL-1999.  
 XX  
 XX 22-JAN-1999; 99WO-FR00137.  
 XX  
 PR 23-JAN-1998; 98FR-0000715.  
 XX  
 PA (INSR ) INST ROUSSY GUSTAVE.  
 XX  
 XX Laurent A, Bellet D;  
 PI  
 XX WPI: 1999-469135/39.  
 DR P-PSDB: AAY26926.  
 XX  
 XX Expression of the INSL4 gene in human embryonic bone tissue and  
 PT ligaments, -  
 XX  
 PS Claim 1; Fig 1; 60pp; French.

XX This sequence represents the human insulin-like 4 gene (INSL4) which  
 CC is expressed in human embryonic bone tissue and ligaments. The 139  
 CC amino acid protein encoded by the gene is designated early-placental  
 CC insulin-like (EPIL) protein. The INSL4 gene encodes 3 different EPIL  
 CC proteins designated EPIL 1, 2 or 3. EPIL 1 is a single chain comprising  
 CC amino acids 18-139, EPIL 2 is a 2 chain protein with chain A comprising  
 CC amino acids 115-139 and chain B comprising amino acids 18-58 and EPIL 3  
 CC is a single chain comprising amino acids 59-114. The nucleic acids and  
 CC protein or antibodies against EPIL 1, 2 or 3, and probes or primers for  
 CC INSL4 are useful for the diagnosis of pathology associated with abnormal  
 CC differentiation and/or proliferation of bone tissue or ligaments or  
 CC development of abnormal cartilage and/or abnormal ossification of forming  
 CC bones, e.g. osteoporosis or dysplasia. Compositions capable of modulation  
 CC differentiation, regeneration and/or proliferation of bone tissue and/or  
 CC ligament cells are also useful in treating bone disease.

XX Sequence 618 BP; 189 A; 145 C; 142 G; 142 T; 0 other;

Query Match 100.0%; Score 342; DB 20; Length 618;  
 Best Local Similarity 100.0%; Pred. No. 1.le-100;  
 Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCAGAGCTGAGGGGATGTGTCGCCGGATTTGGAAACACCTTGTCTCATATTGCCCATG 60  
 DB 182 GCAGAGCTGAGGGGATGTGTCGCCGGATTTGGAAACACCTTGTCTCATATTGCCCATG 241  
 QY 61 CTTGAGAAGACATTCACCAACCCAGGAGGGTGGCTGGAACTCTGGAACTCTGGACGTCCCAAA 120  
 DB 242 CTTGAGAAGACATTCACCAACCCAGGAGGGTGGCTGGAACTCTGGACGTCCCAAA 301  
 QY 121 GAATGTGTCTACCTCCCAACCAAGATGGACAAGCCTTAGTAGTACGACATCAGAATTC 180

```

|||||
302 GAATGGTCAACCTCCACACAAGATGCAAGCCCTTAGTACGACATCAGAAATTC 361
|||||
181 ATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTGAAG 240
|||||
362 ATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTGAAG 421
|||||
241 AAAATATACTTTCCCGCAAAAGAGAACTGGACGTCACAGATTTCATCCATCTGTTGT 300
|||||
422 AAAATATACTTTCCCGCAAAAGAGAACTGGACGTCACAGATTTCATCCATCTGTTGT 481
|||||
301 GAAGTAATTTGTGACGATGAACCTTCAGTTAAATATATGACA 342
|||||
482 GAAGTAATTTGTGACGATGAACCTTCAGTTAAATATATGACA 523

RESULT 5
AAAX27490
ID AAAX27490 standard; DNA; 618 BP.
XX
AC AAAX27490;
XX
DI 22-JUN-1999 (first entry)
XX
DE Human insulin-like gene 4.
XX
KW INSL-4; insulin-like gene; EPIL; early placental insulin-like; antibody;
KW vector; probe; hybridisation; tumour; hypoglycaemia; hyperglycaemia;
KW diabetes; cardiovascular; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 107..526
FT /tag= a
FT /gene= "Insl-4"
FT /product= "EPIL"
FT /note= "early placental insulin-like peptide"

XX
PN WO9909172-A1.
XX
PD 25-FEB-1999.
XX
PF 12-AUG-1998; 98WO-FR01799.
XX
PR 03-NOV-1997; 97FR-0013802.
PR 14-AUG-1997; 97FR-0010387.
XX
PA (INSR ) INST ROUSSY GUSTAVE.
XX
PI Bellet D, Bidart JM, Troalen F, Mock P;
XX
WPI; 1999-181038/15.
XX
P-PSDB; AAW99574.
XX
EPIL polypeptides encoded by insulin-like gene 4 - and corresponding
nucleic acids, antibodies, probes, primers, etc.
XX
Claim 1; Fig 1; 119pp; French.
XX
This sequence represent the INSL-4 (insulin-like gene 4) gene encoding
an Epi (early placental insulin-like) polypeptide. The polypeptide,
antibodies to the polypeptide, vectors containing the coding sequence and
probes derived from the coding sequence, can be used to treat tumours,
preferably angioproliferative tumours, especially Kaposi's sarcoma,
tumours of the pancreas, liver, uterus or breast, angiosarcomas,
glioblastomas, neuroblastomas, rhabdomyosarcomas or leiomyosarcomas;
to promote vascularisation of specific tissues; to treat retinopathy;
macular degeneration, psoriasis, endometriosis, rheumatoid arthritis,
atherosclerosis or hyperthyroidism; to treat post-angioplastic
restenosis; to promote or inhibit embryo implantation; to prevent and/or
treat disorders directly or indirectly connected with insulin-like
activity; to prevent and/or treat disorders directly or indirectly

```

```

CC connected with a dysfunction in carbohydrate metabolism, especially
CC connected with hypo glycaemia or hyperglycaemia, especially gestational
CC diabetes and diabetic complications, especially cardiovascular
CC complications.
XX
SQ Sequence 618 BP; 189 A; 145 C; 142 G; 142 T; 0 other;

Query Match 100.0%; Score 342; DB 20; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.1e-100;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGATGTGGTCCCGGATTTGGAAACACTTGTGTGCATATTGCCCATG 60
|||||
DB 182 GCAGAGCTGAGGGATGTGGTCCCGGATTTGGAAACACTTGTGTGCATATTGCCCATG 241
|||||
QY 61 CCTGAGAGACATTCACACACACCCAGGAGGTGGTGTGGAATCTGGACGTCCCAAA 120
|||||
DB 242 CCTGAGAGACATTCACACACACCCAGGAGGTGGTGTGGAATCTGGACGTCCCAAA 301
|||||
QY 121 GAAATGGTGTCAACCTCCACACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180
|||||
DB 302 GAAATGGTGTCAACCTCCACACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 361
|||||
QY 181 ATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTGAAG 240
|||||
DB 362 ATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTGAAG 421
|||||
QY 241 AAAATATACTTTCCCGCAAAAGAGAGTGGACGTCACAGATTTCATCTGTTGT 300
|||||
DB 422 AAAATATACTTTCCCGCAAAAGAGAGTGGACGTCACAGATTTCATCTGTTGT 481
|||||
QY 301 GAAGTAATTTGTGACGATGAACCTTCAGTTAAATATATGACA 342
|||||
DB 482 GAAGTAATTTGTGACGATGAACCTTCAGTTAAATATATGACA 523
|||||

RESULT 6
AAT68419
ID AAT68419 standard; cDNA; 649 BP.
XX
AC AAT68419;
XX
DT 24-JUL-1997 (first entry)
XX
DE Human relaxin-related factor-2 (RRF-2) cDNA.
XX
KW Relaxin-related factor-2; RRF-2; placenta; growth factor; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 104..523
FT /tag= a
FT sig_peptide 104...181
FT /tag= b
FT mat_peptide 182..520
FT /tag= c
XX
PN WO9716549-A2.
XX
PD 09-MAY-1997.
XX
PF 01-NOV-1996; 96WO-US17342.
XX
PR 21-FEB-1996; 96US-0012016.
PR 03-NOV-1995; 95US-0006221.
XX
PA (REGE-) REGENERON PHARM INC.
XX
PI Davis S;
XX
WPI; 1997-272118/24.
DR P-PSDB; AAW17676.

```

```

XX PT New isolated relaxin-related factor genes - used to develop products
PT PT which can be used in diagnosis and therapy, e.g. in fertility and
PT PT pregnancy applications
XX PS
XX PS Claim 20; Fig 4; 34pp; English.
XX CC A cDNA clone (AA17676), a placenta-specific growth factor related to
CC CC relaxin and to the insulin family of ligands. A search of a
CC CC database of expressed sequence tags revealed the existence of a
CC CC sequence, in human placenta, showing faint homology to relaxin.
CC CC Oligonucleotides corresponding to the database sequence were used
CC CC to probe a human placenta cDNA library. An isolated clone included
CC CC an open reading frame whose predicted protein sequence displayed
CC CC the expected features of an insulin family member and was
CC CC designated RRF-2. A related testis-specific clone, RRF-1 (AA1768418),
CC CC has also been isolated. The isolated clone can be used to produce
CC CC RRF-2 free of other proteins using host-vector systems. RRF-2
CC CC may be useful for modulating the reproductive physiology of mammals
CC CC during pregnancy and parturition.
XX SQ Sequence 649 BP; 204 A; 149 C; 145 G; 151 T; 0 other;

Query Match 100.0%; Score 342; DB 18; Length 649;
Best Local Similarity 100.0%; Pred. No. 1.1e-100;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGATGTGGTCCCGGATTTGGAAACACTTGTCTCATATTGCCCATG 60
Db 179 GCAGAGCTGAGGGATGTGGTCCCGGATTTGGAAACACTTGTCTCATATTGCCCATG 238
QY 61 CCTGAGAGACATTCACACACACCCAGGAGGTGGCTGGATCTGGAGTCCCAAA 120
Db 239 CCTGAGAGACATTCACACACACCCAGGAGGTGGCTGGATCTGGAGTCCCAAA 298
QY 121 GAATGGTGTCACTCCCAACAAAGAGTGGACAGCTTGGTACGACATCAGAAATC 180
Db 299 GAATGGTGTCACTCCCAACAAAGAGTGGACAGCTTGGTACGACATCAGAAATC 358
QY 181 ATTCCTAATTTGCACAGAGTGAAGAACCACTGTCTGAAGGCCAGCCATCATTGA 240
Db 359 ATTCCTAATTTGCACAGAGTGAAGAACCACTGTCTGAAGGCCAGCCATCATTGA 418
QY 241 AAAATAATTAATTTCCCGAAAAAGAGAGTGGAGCTCAGATTTGATCCATTTCTGTGT 300
Db 419 AAAATAATTAATTTCCCGAAAAAGAGAGTGGAGCTCAGATTTGATCCATTTCTGTGT 478
QY 301 GAAGTAATTTGTGACATGGAACCTTCAGTTAAATTAATGATACA 342
Db 479 GAAGTAATTTGTGACATGGAACCTTCAGTTAAATTAATGATACA 520

RESULT 7
ID AAN30186 standard; mRNA; 752 BP.
XX AC
XX AC AAN30186;
XX DT
XX DT 25-MAR-2003 (updated)
XX DT 25-MAY-1992 (first entry)
XX SQ Sequence encoding porcine preprorelaxin.
XX KW Relaxin; hormone; ss.
XX OS Sus scrofa domestica.
XX FH Key Location/Qualifiers
XX FT misc_feature 1..72
XX FT /tag= a
XX FT /label= Claim 7
XX FT misc_feature 73..168

```

```

FT FT /tag= b
FT FT /label= Claim 9
FT FT 73..546
FT FT /tag= c
FT FT /label= Claim 5
FT FT 169..480
FT FT /tag= d
FT FT /label= Claim 10
XX PN EP86649-A.
XX PD 24-AUG-1983.
XX PF 11-FEB-1983; 83EP-0300714.
XX PR 12-FEB-1982; 82AU-0002695.
XX PR 11-FEB-1983; 83AU-0011834.
XX PA (FLOR-) FLOREY INST EXP PHY.
XX PA (FLOR-) FLOREY HOWARD INST.
XX PI Hudson PJ, Haley JD, Niall HD, Shine J;
XX WPI; 1983-748587/35.
XX PT Genes and DNA transfer vectors for prorelaxin expression - useful
XX PT in prodn. of porcine relaxin for veterinary and human use .
XX PS Claim 2; Page 1; 50pp; English.
XX CC The inventors claim synthetic porcine preprorelaxin and prorelaxin
XX CC and synthetic A, B and C peptide chains of prorelaxin, and a gene for
XX CC expression of porcine preprorelaxin or prorelaxin, and their sub-
XX CC units (see AAN30186). They also claim a double-stranded DNA fragment
XX CC for the expression of the signal peptide chain of porcine
XX CC preprorelaxin comprising a coding strand and a complementary strand
XX CC corresp. to a defined mRNA sequence (see AAN30187-N30194) which
XX CC sequences. A probe (AAN30195) is also claimed.
XX CC (Updated on 25-MAR-2003 to correct PF field.)
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 752 BP; 248 A; 143 C; 151 G; 210 U; 0 other;

Query Match 24.8%; Score 84.8; DB 4; Length 752;
Best Local Similarity 49.5%; Pred. No. 3.9e-17;
Matches 107; Conservative 27; Mismatches 82; Indels 0; Gaps 0;

QY 49 TATTGCCCATGCTGTGAGAGACATTCACACACCCAGGAGGTGGTGTGCTGGAATCT 108
Db 136 UGUGGCTCCGUCUCUGGGAAGACUCUCUGGAGAGCCUCGAGCUGGAAACU 195
QY 109 GGACGTCCCAAGAAATGCTCACTCCACACACAAAGATGGACACCTTAGGTAGC 168
Db 196 GGACCCCGGCGAGAACCAUGGCAUCCUCCACCAAGAGUCCAGAAUCCUAAAG 255
QY 169 ACATCAGAAATTCATCTCTAATTTGTCCACAGAGTGAAGAAACCACTGTCTCTGAAGGCAG 228
Db 256 AUGUGGAAUUGUUCUAAUUGCCACAGAGGUGGAGGCAACAACUUGUCUGAGAGGCA 315
QY 229 CCATCATTTGAGAAATATATATCTTTCCGCCCAAAAAG 264
Db 316 CCAUCACUGAGAGGACUACACAAACUACGACUACAAAG 351

RESULT 8
ID AAN30196 standard; mRNA; 752 BP.
XX AC
XX AC AAN30196;
XX DT
XX DT 25-MAR-2003 (updated)
XX DT 25-MAY-1992 (first entry)

```

```

XX DE Sequence encoding porcine preprorelaxin.
XX KW Relaxin; hormone; ss.
XX OS Sus scrofa domestica.
XX FH Key Location/Qualifiers
XX FT CDS 1..549
XX FT /*tag= a
XX PN EP86649-A.
XX PD 24-AUG-1983.
XX PF 11-FEB-1983; 83EP-0300714.
XX PR 12-FEB-1982; 82AU-0002695.
XX PR 11-FEB-1983; 83AU-0011834.
XX PA (FLOR-) FLOREY INST EXP PHY.
XX PA (FLOR-) FLOREY HOWARD INST.
XX PI Hudson PJ, Halsey JD, Niall HD, Shine J;
XX WI; 1983-748587/35.
XX P-PSDB; AAP30392.
XX PT Genes and DNA transfer vectors for preprorelaxin expression - useful
XX PT in prodn. of porcine relaxin for veterinary and human use
XX PS Disclosure; Fig 5; 50pp; English.
XX CC The inventors claim synthetic porcine preprorelaxin and prorelaxin
XX CC and synthetic A, B and C peptide chains of prolaxin, and a gene for
XX CC expression of porcine preprorelaxin or prorelaxin, and their sub-
XX CC units (see AAN30186). They also claim a double-stranded DNA fragment
XX CC for the expression of the signal peptide chain of porcine
XX CC preprorelaxin comprising a coding strand and a complementary strand
XX CC corresp. to a defined mRNA sequence (see AAN30187-N30194) which
XX CC corresponds to the most homologous regions between the pig and rat cDNA
XX CC sequences. A probe (AAN30195) is also claimed.
XX CC (Updated on 25-MAR-2003 to correct PF field.)
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 752 BP; 247 A; 143 C; 151 G; 211 U; 0 other;

Query Match 24.8%; Score 84.8; DB 4; Length 752;
Best Local Similarity 49.5%; Pred. No. 3.9e-17;
Matches 107; Conservative 27; Mismatches 82; Indels 0; Gaps 0;

QY 49 TATTGCCCCATGCTGAGAGAGATTCCACCACCACCCAGGGGTGGTGGTGAATCT 108
DB 136 UGUGGUCGUCGUCGUGGGAAGNACUCUCGAGCAGCCUGAGCUCAGUGGAACU 195

QY 109 GGAGTCCCAAGAAATGGTGTCACTCCACACAAAGATGGACAAAGCCTTAGGTACG 168
DB 196 GGAGCCCGGCGAAGAACCAUGCCAUCCAUCCAAAGACGAGAAAUUAAAGAU 255

QY 169 ACATCAGATTTCCTTAATTTGTCACAGAGCTGAAGAAACCTGCTGAAGGCAG 228
DB 256 AUGUUGAAGUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU 315

QY 229 CCATCATTTGAAGAAATTAATATTCTTCCCGCAAAAG 264
DB 316 CCAUCAGAGAGAGAGCUCACACACAUUCGCAUACAG 351

RESULT 9
AAS69640/c
ID AAS69640 standard; cDNA; 2364 BP.
XX
AC AAS69640;

```

---

```

XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #5444.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX WI; 2001-639362/73.
XX P-PSDB; ABG05453.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 1; SEQ ID No 5444; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC (II). (II) is useful for generating antibodies against it, detecting or
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 2364 BP; 861 A; 537 C; 439 G; 527 T; 0 other;

Query Match 23.8%; Score 81.4; DB 23; Length 2364;
Best Local Similarity 74.1%; Pred. No. 8.1e-16;
Matches 103; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 117 CAAGAAGATGTGTCAACCTCCACACAAAGATGGACAGCCTTAGGTAGCATCAGA 176
DB 351 CACAGAGATGATGTATCATCCACCCTAATAAGATGCAGGACCTTAGATATGAGTTGG 292

QY 177 ATTCAATCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCATCAT 236
DB 291 ATTCAATCTTAATTTGTCACAGAGCTGAAGAAACCAATTTCTGTGAGGGGTAGCTGCT 232

QY 237 GAAGAAATAATACTTTCC 255
DB 231 GAGAAAGTACTACTTTCC 213

```

[illegible]

[illegible]

RESULT 12  
AAQ06595  
ID AAQ06595 standard; DNA: 489 BP.  
XX  
XX AC AAQ06595;  
XX AC  
XX DT 27-FEB-1991 (first entry)  
XX DT  
DE H2 proteolaxin gene insert in plasmid pTRProRelasp.  
XX

KW Relaxin; PR; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9013659-A.  
 XX  
 PD 15-NOV-1990.  
 XX  
 PF 16-APR-1990; 90WO-US02085.  
 XX  
 PR 04-MAY-1989; 89DS-0347550.  
 XX  
 XX (GETH ) GENENTECH INC.  
 XX  
 PA Henner DJ, Vandlen RI, Wikins JA, Vansura D;  
 PI  
 XX WPI; 1990-361494/48.  
 DR P-PSDB; AAR07987.  
 XX  
 XX  
 XX  
 PT  
 PT treating free-cysteine form of polypeptide at desired junction.  
 XX  
 PS Disclosure: Fig 2A; 56pp; English.

The plasmid was prep'd. from a clone isolated from a cDNA library  
prep'd. from RNA isolated from human corpus luteum, screened with an  
H1-cDNA probe. A fragment encoding the N-terminal of PR was iso-  
lated and was ligated with a 410 bp fragment encoding AAS 17-153 of  
PR, and a fragment from pGHG2071+L (including the Tsp promoter, AAS  
1-137 of methGHG, and amp and tet resistance. A portion of the  
resulting construction (pPRproH2) was ligated with a fragment from  
the original clone encoding the PR C-terminal, and a fragment from  
pBR322XAP encoding part of the beta-lactamase gene. This produced  
plasmid pPRproREL. A SstI signal sequence was then fused to the  
PR gene and a portion including the gene and signal sequence was  
ligated into a vector identical to pPRSTIGH in which the HGH  
gene had been removed. The resulting construction, pPRSTIIpRoREL  
was treated to remove the SstII sequence and the first 11 AAs of H2  
PR. This was replaced with a synthetic DNA duplex encoding the  
first 12 AAs of H2 PR (including Asp1), to produce pPRproRELasp,  
the coding part of which is shown below. The plasmid can be used  
to construct vectors that encode Asp-inserted H2PR with enhanced  
acid cleavage sites.  
See also AAQ06596.

Sequence 489 BP; 156 A; 104 C; 103 G; 126 T; 0 other;

```
Query Match      21.0%; Score 71.8; DB 11; Length 489;
Best Local Similarity 66.5%; Pred. NO. 5.3e-13;
Matches 103; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
```

[illegible]

142	QY	AACAAAGATGACACAGCCTT	AGTACGACATCAGAA	TTCAATTCCTAA	TTTGTCCACAGAG	201
163	Db	AAACAAAGATACAGAAACCAT	AAATATATGATGTCAGAA	TTTGTGCTAA	TTTGGCCACAGGAG	222
202	QY	CTGAAGAAACCACTGCT	CTGAAGGGCAGCCAT	CA	TT	236
223	Db	CTGAAGTTAA	CCCTGCTGAGATGAC	CCAGCA	TT	257

RESULT 13  
AAQ06596  
ID AAQ06596 standard; DNA; 495 BP.  
XX  
XX  
AC  
AC  
AAQ06596;  
XX  
XX  
DT  
DT  
27-FEB-1991 (first entry)  
XX  
XX  
DE  
DE  
H2 proralexin gene insert in plasmid pTR411.

KW Relaxin; PR; ds.

XX Homo sapiens.

XX PN W09013659-A

XX  
PD  
15-NOV-1990.XX  
PF  
16-APR-1990:XX  
PR  
04-MAY-1989:XX  
PA (GETH) GENEXX  
PT  
Hennepin DTXX  
DP  
WPT: 1990-36

DR P-PSDB; AARO  
XX

PT Cleaving pol  
treating fre  
PT

XX  
DS  
Discharge:

2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840. 841.

CC pTtpProRelAs

CC PR gene betw

CC introduce an

CC regions. Pla

[illegible]

CC PR):

CC addnl. Asp c

CC an extra ASP

CC acid cleavage

XX  
XX  
OCT 20 1966

sequence 495

Query Match  
Best Local Similarity

**Matches** 103;

Qy 82 ACC 11

**Db** 106 AGC



|    |     |                                        |
|----|-----|----------------------------------------|
| QY | 142 | ACCAAGATGGACAAGCCTTAGTACGACATCAGAAATTC |
|    | 4   |                                        |
| DB | 166 | ACAAGATACAGAAACCATAATATGATGTCAGAAATTT  |
|    |     |                                        |
| QY | 202 | CTGAAGAACCACTGTCTGAAGGCCAGCCATCATT     |
|    |     |                                        |
| DB | 226 | CTGAAGTTACCTGTCTCAGATGACGCCAGCAT       |
|    |     |                                        |

|           |                                                   |
|-----------|---------------------------------------------------|
| RESULT 14 |                                                   |
| AAN40086  |                                                   |
| ID        | AAN40086 standard; mRNA; 558 BP.                  |
| XX        |                                                   |
| XX        |                                                   |
| AC        | AAN40086;                                         |
| XX        |                                                   |
| XX        |                                                   |
| DT        | 25-MAR-2003 (updated)                             |
| DT        | 04-FEB-1992 (first entry)                         |
| XX        |                                                   |
| XX        |                                                   |
| DE        | Sequence encoding human preprorelaxin H2.         |
| XX        |                                                   |
| XX        |                                                   |
| KW        | Relaxin; hormone; probe; uterine contraction; ss. |
| XX        |                                                   |
| OS        | H2.                                               |

Query Match 21.0%; Score 71.8; DB 5; Length 558;  
Best Local Similarity 52.3%; Pred. NO. 5.6e-13;  
Matches 81; Conservative 22; Mismatches 52; Indels 0; Gaps 0;

|           |                                                                           |                                                                |     |
|-----------|---------------------------------------------------------------------------|----------------------------------------------------------------|-----|
| Qy        | 82                                                                        | ACCCAGGAGGTGCTGCTGGAATCTTGAGAGTCCCAAGAAATGGTGTCAACCTCCAA       | 141 |
| Db        | 172                                                                       | AGCCAGGAAGUGCCUCUCAGAACCCUAGACCAGGCGCAGAAAUUGGCCAUCCUCAUC      | 231 |
| Qy        | 142                                                                       | AAACAAGATGGAGAAGCCCTTAGGTACGACATCAGAATTCATTTCTTAATTTGTCAACAGAG | 201 |
| Db        | 232                                                                       | ACAAGAATACAGAACCAACCAUAAATGAGUGUCAGAAUUGUGCUAAUUGCCACAGGAG     | 291 |
| Qy        | 202                                                                       | CTGAGAGAACCACACTGCTCTGAAGGGCGCCATCATTT                         | 236 |
| Db        | 292                                                                       | CUGAAGUUAACCCUGUCUGAUGCAGCCAGCAU                               | 326 |
| RESULT 15 |                                                                           |                                                                |     |
| AA        | AN92483                                                                   |                                                                |     |
| ID        | AAAN92483 standard; mRNA; 558 BP.                                         |                                                                |     |
| XX        | AAAN92483;                                                                |                                                                |     |
| XX        | 25-MAR-2003 (updated)                                                     |                                                                |     |
| DT        | 21-JUN-1990 (first entry)                                                 |                                                                |     |
| XX        | XX                                                                        |                                                                |     |
| XX        | mRNA sequence of human preprorelaxin H2.                                  |                                                                |     |
| DE        | Relaxin; H2-relaxin; pubic symphysis.                                     |                                                                |     |
| XX        | Homo sapiens.                                                             |                                                                |     |
| OS        | EP303033-A.                                                               |                                                                |     |
| PN        | 15-FEB-1989.                                                              |                                                                |     |
| XX        | 12-DEC-1983; 88EP-0110103.                                                |                                                                |     |
| XX        | 13-DEC-1982; 82AU-0007247.                                                |                                                                |     |
| XX        | (FLOR-) FLOREY INST EXPERIMENTAL PHYSIOLOGY.                              |                                                                |     |
| PA        | Hudson PJ, Tregear GW, Niall HD;                                          |                                                                |     |
| XX        | WPI; 1989-047874/07.                                                      |                                                                |     |
| XX        | P-PSDB; AAP94621.                                                         |                                                                |     |
| XX        | New human H2-relaxin analogues -                                          |                                                                |     |
| PT        | with shortened and/or modified A and/or B chains.                         |                                                                |     |
| PT        | Disclosure; Page ?; 25pp; English.                                        |                                                                |     |
| XX        | XX                                                                        |                                                                |     |
| XX        | H2 relaxin is deduced from a cDNA clone, it has the general properties of |                                                                |     |
| CC        | a growth factor and is capable of altering nature of connective tissue    |                                                                |     |
| CC        | and inducing smooth muscle contraction, specifically during labour.       |                                                                |     |
| CC        | (Updated on 25-MAR-2003 to correct PF field.)                             |                                                                |     |
| CC        | (Updated on 25-MAR-2003 to correct PA field.)                             |                                                                |     |
| CC        | (Updated on 25-MAR-2003 to correct PI field.)                             |                                                                |     |
| XX        | Sequence 558 BP; 169 A; 122 C; 116 G; 151 U; 0 other;                     |                                                                |     |

| Query Match           | 21.0%           | Score 71.8                                                    | DB 10    | Length 558 |
|-----------------------|-----------------|---------------------------------------------------------------|----------|------------|
| Best Local Similarity | 52.3%           | Pred. No. 5.6e-13                                             |          |            |
| Matches 81            | Conservative 22 | Mismatches 52                                                 | Indels 0 | Gaps 0     |
| Qy                    | 82              | ACCCACGAGGGTGGCTGCTGGAAATCTGGAGTCCCAAGAGAAATGGTCTCAACCTCCAC   | 141      |            |
| Db                    | 172             | AGCCAGGAUGUCCUCAGACACCCAGACAGCAGUGGCAGAAAUUGGCCAUCUCCAU       | 231      |            |
| Qy                    | 142             | AACAAGATGGACAAGCGCTTAGGTACGACATCAGAAATTCATTCCTAATTTGTCACACGAG | 201      |            |
| Db                    | 232             | AACAAGAUACAGAAACCAUAAUAUGAUGUCACAAUUGUGCUAAUUGCCACAGAG        | 291      |            |
| Qy                    | 202             | CTGAAGAAACACTGTCTGAAGGGCAGCCCATTT                             | 236      |            |
| Db                    | 292             | CUGAAGUUAACCCUGUGUGAUGAUGACGCCAGCAU                           | 326      |            |

|    |     |                                    |     |
|----|-----|------------------------------------|-----|
| QY | 202 | CTGAAGAAACCACTCTCTGAAGGCGCCATCATT  | 236 |
|    |     | :   :   :                 : :      |     |
| Db | 292 | CUGAAGUUAACCCGUCUCUGAUGCAGCCAGCAUU | 326 |

Search completed: October 17, 2003, 01:52:32  
Job time : 207 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2003, 01:37:56 ; Search time 52 Seconds  
(without alignments)  
2902.943 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 342

Sequence: 1 GCAGAGCTGAGGAGTGG.....CTTCACTTAATATGTACA 342

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Issued\_Patents\_NA.\*
- 2: /cgn2.6/ptodata/1/ina/5A.COMB.seq.\*
- 3: /cgn2.6/ptodata/1/ina/5B.COMB.seq.\*
- 4: /cgn2.6/ptodata/1/ina/6A.COMB.seq.\*
- 5: /cgn2.6/ptodata/1/ina/6B.COMB.seq.\*
- 6: /cgn2.6/ptodata/1/ina/PCTUS.COMB.seq.\*
- 7: /cgn2.6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 342   | 100.0       | 420    | 3  | US-08-991-890-1    |
| 2          | 342   | 100.0       | 480    | 3  | US-08-991-890-3    |
| 3          | 342   | 100.0       | 615    | 2  | US-08-482-842B-3   |
| 4          | 342   | 100.0       | 934    | 3  | US-09-174-465D-4   |
| 5          | 342   | 100.0       | 934    | 4  | US-09-599-564A-4   |
| 6          | 340.4 | 99.5        | 615    | 2  | US-08-482-842B-1   |
| 7          | 340.4 | 99.5        | 615    | 3  | US-09-174-465D-1   |
| 8          | 340.4 | 99.5        | 615    | 4  | US-09-599-564A-1   |
| 9          | 153   | 44.7        | 153    | 3  | US-09-174-465D-9   |
| 10         | 153   | 44.7        | 153    | 4  | US-09-599-564A-9   |
| 11         | 99    | 28.9        | 123    | 3  | US-09-174-465D-7   |
| 12         | 99    | 28.9        | 123    | 4  | US-09-599-564A-7   |
| 13         | 90    | 26.3        | 93     | 3  | US-09-174-465D-11  |
| 14         | 90    | 26.3        | 93     | 4  | US-09-599-564A-11  |
| 15         | 71.8  | 21.0        | 489    | 6  | 5464756-17         |
| 16         | 71.8  | 21.0        | 495    | 6  | 5464756-19         |
| 17         | 71.8  | 21.0        | 915    | 1  | US-08-443-568B-11  |
| 18         | 71.8  | 21.0        | 915    | 5  | PCT-US94-06997-11  |
| c 19       | 69    | 20.2        | 231    | 4  | US-09-439-313-465  |
| c 20       | 69    | 20.2        | 231    | 4  | US-09-352-616A-465 |
| c 21       | 38.6  | 11.3        | 2574   | 3  | US-09-142-529-2    |
| c 22       | 38.6  | 11.3        | 2574   | 4  | US-10-045-428A-2   |
| c 23       | 35    | 10.2        | 116592 | 4  | US-09-818-512-3    |
| c 24       | 34.4  | 10.1        | 1001   | 4  | US-09-641-638-123  |
| c 25       | 31    | 9.1         | 2585   | 3  | US-09-008-697A-7   |
| c 26       | 30.2  | 8.8         | 2895   | 4  | US-09-422-936-52   |
| c 27       | 30    | 8.8         | 932    | 4  | US-09-257-179-20   |

|      |      |     |      |   |                    |                   |
|------|------|-----|------|---|--------------------|-------------------|
| c 28 | 30   | 8.8 | 2250 | 1 | US-08-483-639-1    | Sequence 1, Appl  |
| c 29 | 30   | 8.8 | 5093 | 1 | US-08-468-036-23   | Sequence 23, Appl |
| c 30 | 30   | 8.8 | 5093 | 2 | US-08-376-843-23   | Sequence 23, Appl |
| c 31 | 29.6 | 8.7 | 9069 | 4 | US-08-961-527-37   | Sequence 37, Appl |
| c 32 | 29.4 | 8.6 | 553  | 2 | US-08-721-488-7    | Sequence 7, Appl  |
| c 33 | 29.4 | 8.6 | 1789 | 4 | US-08-936-165A-116 | Sequence 116, App |
| c 34 | 29.4 | 8.6 | 1875 | 4 | US-09-511-881A-11  | Sequence 11, Appl |
| c 35 | 29.4 | 8.6 | 2451 | 2 | US-08-820-170A-29  | Sequence 29, Appl |
| c 36 | 29.4 | 8.6 | 2451 | 3 | US-09-055-699-29   | Sequence 29, Appl |
| c 37 | 29.4 | 8.6 | 2451 | 3 | US-09-273-565-29   | Sequence 29, Appl |
| c 38 | 29.4 | 8.6 | 2451 | 4 | US-09-565-538-29   | Sequence 29, Appl |
| c 39 | 29.4 | 8.6 | 2451 | 4 | US-09-661-468-29   | Sequence 29, Appl |
| c 40 | 29.4 | 8.6 | 2451 | 4 | US-09-976-155-29   | Sequence 29, Appl |
| c 41 | 29.4 | 8.6 | 2487 | 2 | US-08-820-170A-32  | Sequence 32, Appl |
| c 42 | 29.4 | 8.6 | 2487 | 3 | US-09-055-699-32   | Sequence 32, Appl |
| c 43 | 29.4 | 8.6 | 2487 | 3 | US-09-273-565-32   | Sequence 32, Appl |
| c 44 | 29.4 | 8.6 | 2487 | 4 | US-09-565-538-32   | Sequence 32, Appl |
| c 45 | 29.4 | 8.6 | 2487 | 4 | US-09-661-468-32   | Sequence 32, Appl |

ALIGNMENTS

RESULT 1  
US-08-991-890-1  
; Sequence 1, Application US/08991890  
; Patent No. 6114307  
; GENERAL INFORMATION:  
; APPLICANT: Jaspers, Stephen R.  
; APPLICANT: Sprugel, Katherine H.  
; APPLICANT: Ren, Hong Ping  
; APPLICANT: Humes, Jacqueline M.  
; APPLICANT: Hoffman, Ross C.  
; APPLICANT: Conklin, Dorell C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; NUMBER OF SEQUENCES: 7  
; TITLE OF INVENTION: STIMULATING PANCREATIC ISLET CELL REGENERATION  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ZymoGenetics, Inc.  
; STREET: 1201 Eastlake Avenue East  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/991.890  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/033.003  
; FILING DATE: December 16, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sawislak, Deborah A.  
; REGISTRATION NUMBER: 37,438  
; REFERENCE/DOCKET NUMBER: 96-41  
; TELEPHONE: 206-442-6672  
; TELEFAX: 206-442-6678  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 420 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: Coding Sequence

```
; LOCATION: 1...417
; OTHER INFORMATION:
US-08-991-890-1

Query Match      100.0%  Score 342; DB 3; Length 420;
Best Local Similarity 100.0%  Pred. No. 1.7e-101;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGGTCCCGATTGGAAAAACACTTGTGTCATATTGCCCATG 60
   |||||||
DB 76 GCAGAGCTGAGGGGATGTGGTCCCGATTGGAAAAACACTTGTGTCATATTGCCCATG 135
   |||||||
QY 61 CCTGAGAAGACATTCACCAACACCCAGAGGGTGGTCTGGAATCTGGACGTCCCAAA 120
   |||||||
DB 136 CCTGAGAAGACATTCACCAACACCCAGAGGGTGGTCTGGAATCTGGACGTCCCAAA 195
   |||||||
QY 121 GAAATGGTGTCAACCTCCACACAAACAGATGGCAAGCCATTAGTAGACATCAGATTC 180
   |||||||
DB 196 GAAATGGTGTCAACCTCCACACAAACAGATGGCAAGCCATTAGTAGACATCAGATTC 255
   |||||||
QY 181 ATTCTTAATTCTCACCAGAGCTGAAGAAACCACTGTCTGAAGGCGAGCCATCATTGAAG 240
   |||||||
DB 256 ATTCTTAATTCTCACCAGAGCTGAAGAAACCACTGTCTGAAGGCGAGCCATCATTGAAG 315
   |||||||
QY 241 AAAATAATACTTTCCCGCAAAAGAGAGTGGAGCTACAGATTGTATCCATCTGTGT 300
   |||||||
DB 316 AAAATAATACTTTCCCGCAAAAGAGAGTGGAGCTACAGATTGTATCCATCTGTGT 375
   |||||||
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 342
   |||||||
DB 376 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 417
   |||||||

RESULT 2
US-08-991-890-3
; Sequence 3, Application US/08991890
; Patent No. 6114307
; GENERAL INFORMATION:
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Sprugel, Katherine H.
; APPLICANT: Rep, Hong Ping
; APPLICANT: Humes, Jacqueline M.
; APPLICANT: Hoffman, Ross C.
; APPLICANT: Conklin, Darrell C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: STIMULATING PANCREATIC ISLET CELL REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,890
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,003
; FILING DATE: December 16, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 96-41
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678

; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...477
; OTHER INFORMATION:
US-08-991-890-3

Query Match      100.0%  Score 342; DB 3; Length 480;
Best Local Similarity 100.0%  Pred. No. 1.8e-101;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGGTCCCGATTGGAAAAACACTTGTCTGTCTATATTGCCCATG 60
   |||||||
DB 136 GCAGAGCTGAGGGGATGTGGTCCCGATTGGAAAAACACTTGTCTGTCTATATTGCCCATG 195
   |||||||
QY 61 CCTGAGAAGACATTCACCAACACCCAGAGGGTGGTCTGGAATCTGGACGTCCCAAA 120
   |||||||
DB 196 CCTGAGAAGACATTCACCAACACCCAGAGGGTGGTCTGGAATCTGGACGTCCCAAA 255
   |||||||
QY 121 GAAATGGTGTCAACCTCCACACAAACAGATGGCAAGCCATTAGTAGACATCAGATTC 180
   |||||||
DB 256 GAAATGGTGTCAACCTCCACACAAACAGATGGCAAGCCATTAGTAGACATCAGATTC 315
   |||||||
QY 181 ATTCTTAATTCTCACCAGAGCTGAAGAAACCACTGTCTGAAGGCGAGCCATCATTGAAG 240
   |||||||
DB 316 ATTCTTAATTCTCACCAGAGCTGAAGAAACCACTGTCTGAAGGCGAGCCATCATTGAAG 375
   |||||||
QY 241 AAAATAATACTTTCCCGCAAAAGAGAGTGGAGCTACAGATTGTATCCATCTGTGT 300
   |||||||
DB 376 AAAATAATACTTTCCCGCAAAAGAGAGTGGAGCTACAGATTGTATCCATCTGTGT 435
   |||||||
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 342
   |||||||
DB 436 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 477
   |||||||

RESULT 3
US-08-482-842B-3
; Sequence 3, Application US/08482842B
; Patent No. 5910480
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahmet
; APPLICANT: BELLET, Dominique
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN,
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF THIS PROTEIN AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION CONTAINING SUCH, DNA CODING FOR
; TITLE OF INVENTION: SAID PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Svecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,842B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: FR 94-07191
; FILING DATE: 13-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 017753-058
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..523
; DS-08-482-842B-3

Query Match      100.0%; Score 342; DB 2; Length 615;
Best Local Similarity 100.0%; Pred. No. 2e-101;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACTTGTGCTGCATATTTGCCCCCATG 60
DB 182 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACTTGTGCTGCATATTTGCCCCCATG 241
QY 61 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGCTGGAATCTGGAGCTGCCCAAA 120
DB 242 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGCTGGAATCTGGAGCTGCCCAAA 301
QY 121 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180
DB 302 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 361
QY 181 ATTCCTAATTTGTCACCAAGCTGGAAGAACCCACTGCTGAAGGCGACCCATCATTTGAAG 240
DB 362 ATTCCTAATTTGTCACCAAGCTGGAAGAACCCACTGCTGAAGGCGACCCATCATTTGAAG 421
QY 241 AAAATAATACTTTCCCGGAAAAGAGAGTGGACGTACAGATTTGATTCATTTCTGTGT 300
DB 422 AAAATAATACTTTCCCGGAAAAGAGAGTGGACGTACAGATTTGATTCATTTCTGTGT 481
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 342
DB 482 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 523

RESULT 4
US-09-174-465D-4
; Sequence 4, Application US/09174465D
; Patent No. 6180364
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahment
; APPLICANT: CHASSIN, Dominique
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; FILE REFERENCE: 017753-103
; CURRENT APPLICATION NUMBER: US/09/174,465D
; CURRENT FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 934
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
; DS-09-599-564A-4

Query Match      100.0%; Score 342; DB 4; Length 934;
Best Local Similarity 100.0%; Pred. No. 2.4e-101;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAACAACTTGTGCTGCATATTTGCCCCCATG 60
DB 501 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAACAACTTGTGCTGCATATTTGCCCCCATG 560
QY 61 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGCTGGAATCTGGAGCTGCCCAAA 120
DB 501 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAACAACTTGTGCTGCATATTTGCCCCCATG 560
QY 61 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGCTGGAATCTGGAGCTGCCCAAA 120
```

```
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
; DS-09-174-465D-4

Query Match      100.0%; Score 342; DB 3; Length 934;
Best Local Similarity 100.0%; Pred. No. 2.4e-101;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAACAACTTGTGCTGCATATTTGCCCCCATG 60
DB 501 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAACAACTTGTGCTGCATATTTGCCCCCATG 560
QY 61 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGCTGGAATCTGGAGCTGCCCAAA 120
DB 561 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGCTGGAATCTGGAGCTGCCCAAA 620
QY 121 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180
DB 621 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 680
QY 181 ATTCCTAATTTGTCACCAAGCTGGAAGAACCCACTGCTGAAGGCGACCCATCATTTGAAG 240
DB 681 ATTCCTAATTTGTCACCAAGCTGGAAGAACCCACTGCTGAAGGCGACCCATCATTTGAAG 740
QY 241 AAAATAATACTTTCCCGGAAAAGAGAGTGGACGTACAGATTTGATTCATTTCTGTGT 300
DB 741 AAAATAATACTTTCCCGGAAAAGAGAGTGGACGTACAGATTTGATTCATTTCTGTGT 800
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 342
DB 801 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 842

RESULT 5
US-09-599-564A-4
; Sequence 4, Application US/09599564A
; Patent No. 6362318
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahment
; APPLICANT: CHASSIN, Dominique
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; FILE REFERENCE: 017753-127
; CURRENT APPLICATION NUMBER: US/09/599,564A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/174,465
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: US 08/482,842
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 934
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
; DS-09-599-564A-4

Query Match      100.0%; Score 342; DB 4; Length 934;
Best Local Similarity 100.0%; Pred. No. 2.4e-101;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAACAACTTGTGCTGCATATTTGCCCCCATG 60
DB 501 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAACAACTTGTGCTGCATATTTGCCCCCATG 560
QY 61 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGCTGGAATCTGGAGATCTGGACCTGCCAAA 120
```

```
Db 561 CCTGAGAGACATTCACCACCACCCAGAGGCTGCTGGAATCTGGAGTCCCAAA 620
Qy 121 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 180
Db 621 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 680
Qy 181 ATTCCTAATTTGTACACAGAGCTGAAGAAACCACTGCTCTGAAGGGCAGCCATCATTGAAG 240
Db 681 ATTCCTAATTTGTACACAGAGCTGAAGAAACCACTGCTCTGAAGGGCAGCCATCATTGAAG 740
Qy 241 AAATAATATCTTTCCCGAAAAAGAGAGTGGAGCTCAGATTTGATCCATCTGTTGT 300
Db 741 AAATAATATCTTTCCCGAAAAAGAGAGTGGAGCTCAGATTTGATCCATCTGTTGT 800
Qy 301 GAAGTAATTTGTGAGGTGAACCTTCAGTTAAATATGTACA 342
Db 801 GAAGTAATTTGTGAGGTGAACCTTCAGTTAAATATGTACA 842

RESULT 6
US-08-482-842B-1
; Sequence 1, Application US/08482842B
; Patent No. 5910480
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahmet
; APPLICANT: CHASSIN, Dordine
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN,
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF THIS PROTEIN AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION CONTAINING SUCH, DNA CODING FOR
; TITLE OF INVENTION: SAID PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Buras, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482.842B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94-07191
; FILING DATE: 13-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 017753-058
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..615
; US-08-482-842B-1

Query Match 99.5%; Score 340.4; DB 2: Length 615;
Best Local Similarity 99.7%; Pred. No. 6.6e-101;
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 561 CCTGAGAGACATTCACCACCACCCAGAGGCTGCTGGAATCTGGAGTCCCAAA 620
Qy 121 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 180
Db 621 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 680
Qy 181 ATTCCTAATTTGTACACAGAGCTGAAGAAACCACTGCTCTGAAGGGCAGCCATCATTGAAG 240
Db 681 ATTCCTAATTTGTACACAGAGCTGAAGAAACCACTGCTCTGAAGGGCAGCCATCATTGAAG 740
Qy 241 AAATAATATCTTTCCCGAAAAAGAGAGTGGAGCTCAGATTTGATCCATCTGTTGT 300
Db 741 AAATAATATCTTTCCCGAAAAAGAGAGTGGAGCTCAGATTTGATCCATCTGTTGT 800
Qy 301 GAAGTAATTTGTGAGGTGAACCTTCAGTTAAATATGTACA 342
Db 801 GAAGTAATTTGTGAGGTGAACCTTCAGTTAAATATGTACA 842
```

```
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAAAACACTTGTCTGTCATATTGCCCATG 60
Db 181 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAAAACACTTGTCTGTCATATTGCCCATG 240
Qy 61 CCTGAGAAGACATTCACCACCACCCAGAGGCTGGCTGTCTGGAATCTGGACGTCCCAAA 120
Db 241 CCTGAGAAGACATTCACCACCACCCAGAGGCTGGCTGTCTGGAATCTGGACGTCCCAAA 300
Qy 121 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 180
Db 301 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 360
Qy 181 ATTCCTAATTTGTACACAGAGCTGAAGAAACCACTGCTCTGAAGGGCAGCCATCATTGAAG 240
Db 361 ATTCCTAATTTGTACACAGAGCTGAAGAAACCACTGCTCTGAAGGGCAGCCATCATTGAAG 420
Qy 241 AAATAATATCTTTCCCGAAAAAGAGAGTGGAGCTCAGATTTGATCCATCTGTTGT 300
Db 421 AAATAATATCTTTCCCGAAAAAGAGAGTGGAGCTCAGATTTGATCCATCTGTTGT 480
Qy 301 GAAGTAATTTGTGAGGTGAACCTTCAGTTAAATATGTACA 342
Db 481 GAAGTAATTTGTGAGGTGAACCTTCAGTTAAATATGTACA 522

RESULT 7
US-09-174-465D-1
; Sequence 1, Application US/09174465D
; Patent No. 6180364
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahmet
; APPLICANT: CHASSIN, Dordine
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 017753-103
; CURRENT APPLICATION NUMBER: US/09/174,465D
; MEDIUM TYPE: Floppy disk
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: US 08/482,842
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(615)
; OTHER INFORMATION: Description of Unknown Organism: EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like Peptide
; US-09-174-465D-1

Query Match 99.5%; Score 340.4; DB 3: Length 615;
Best Local Similarity 99.7%; Pred. No. 6.6e-101;
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAAAACACTTGTCTGTCATATTGCCCATG 60
Db 181 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAAAACACTTGTCTGTCATATTGCCCATG 240
Qy 61 CCTGAGAAGACATTCACCACCACCCAGAGGCTGGCTGTCTGGAATCTGGACGTCCCAAA 120
Db 241 CCTGAGAAGACATTCACCACCACCCAGAGGCTGGCTGTCTGGAATCTGGACGTCCCAAA 300
Qy 121 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 180
Db 301 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 360
```

| Query Match           | 99.5%;          | Score 340.4;                                                   | DB 4;     | Length 615; |
|-----------------------|-----------------|----------------------------------------------------------------|-----------|-------------|
| Best Local Similarity | 99.7%;          | Pred. No. 6.6e-101;                                            |           |             |
| Matches 341;          | Conservative 0; | Mismatches 1;                                                  | Indels 0; | Gaps 0      |
| Qy                    | 1               | GCAGAGCTGAGGGGATGTGGTCCCCGATTTGGAAACACATTGCTGTATATTCGCCCATG    | 60        |             |
| Db                    | 181             | GCAGAGCTGAGGGGATGTGGTCCCCGATTTGGAAACACATTGCTGTATATTCGCCCATG    | 240       |             |
| Qy                    | 61              | CCTGAGAAGACATTCAACCAACCACCCAGGAGGGTGGTGTGGAATCTGGACGCTCCCAA    | 120       |             |
| Db                    | 241             | CCTGAGAAGACATTCAACCAACCACCCAGGAGGGTGGTGTGGAATCTGGACGCTCCCAA    | 300       |             |
| Qy                    | 121             | GAATGGGTCAACTCCCAACCAACAGATGGACAGCCCTTAGTACGACATCAGAAATTC      | 180       |             |
| Db                    | 301             | GAATGGGTCAACTCCCAACCAACAGATGGACAGCCCTTAGTACGACATCAGAAATTC      | 360       |             |
| Qy                    | 181             | ATTCCCTAAATTTGTCACGAGCTGAAGAAACCACCTGTCTGAAGGCGACCATCATTTGAAG  | 240       |             |
| Db                    | 361             | ATTCCCTAAATTTGTCACGAGCTGAAGAAACCACCTGTCTGAAGGCGACCATCATTTGAAG  | 420       |             |
| Qy                    | 241             | AAAATAATTAATTTCCCGCAAAAAGAGAAGTGGACGCTCAGAGATTTGATTCATTCTGTTGT | 300       |             |
| Db                    | 421             | AAAATAATTAATTTCCCGCAAAAAGAGAAGTGGACGCTCAGAGATTTGATTCATTCTGTTGT | 480       |             |
| Qy                    | 301             | GAAGTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA                    | 342       |             |
| Db                    | 481             | GAAGTAATTTGTGACGATGGAACTTCAGTTAAATTTATGTACA                    | 522       |             |

RESULT 10  
US-09-599-564A-9  
; Sequence 9, Application US/09599564A  
; Patent No. 6362318  
; GENERAL INFORMATION:  
; APPLICANT: KOMAN, Aiment  
; APPLICANT: CHASSIN, Dorline  
; APPLICANT: BELLET, Dominique  
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIPLACENTIN, PROCESS FOR THE  
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL  
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID  
; TITLE OF INVENTION: PROTEIN  
; FILE REFERENCE: 017753-127  
; CURRENT APPLICATION NUMBER: US/09/599,564A  
; CURRENT FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 09/174,465  
; PRIOR FILING DATE: 1998-10-19  
; PRIOR APPLICATION NUMBER: US 08/482,842  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 153  
; TYPE: DNA  
; ORGANISM: Unknown

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(153)
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
US-09-599-564A-9

Query Match          44.78; Score 153; DB 4; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.2e-40;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 CTGGAATCTGACGCTCCCAAGAAATGGTGTCAACCTCCCAACAAAGATGCAAGCC 159
      |||||||
Db 1 CTGGAATCTGACGCTCCCAAGAAATGGTGTCAACCTCCCAACAAAGATGCAAGCC 60

Qy 160 TTAGGTAGCAGATCAGAAATTCATTCCTAAATTTGTCAACGAGCTGAAGAAACCACTGTCT 219
      |||||||
Db 61 TTAGGTAGCAGATCAGAAATTCATTCCTAAATTTGTCAACGAGCTGAAGAAACCACTGTCT 120

Qy 220 GAAGGGCAGCCATCATTTGAAGAAATAATACTT 252
      |||||||
Db 121 GAAGGGCAGCCATCATTTGAAGAAATAATACTT 153

RESULT 11
US-09-174-465D-7
; Sequence 7, Application US/09174465D
; Patent No. 6180364
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahment
; APPLICANT: CHASSIN, Dorine
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 017753-103
; CURRENT APPLICATION NUMBER: US/09/174,465D
; PRIOR FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 123
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(123)
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
US-09-174-465D-7

Query Match          28.9%; Score 99; DB 3; Length 123;
Best Local Similarity 100.0%; Pred. No. 6.5e-23;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCAGAGCTGAGGGATGTGTCCTCCGATTTGGAACACACTTGTCTCATATTGCCCATG 60
      |||||||
Db 25 GCAGAGCTGAGGGATGTGTCCTCCGATTTGGAACACACTTGTCTCATATTGCCCATG 84

Qy 61 CCTGAGAAGACATTCACCAACCCAGGAGGGTGGCTG 99
      |||||||
Db 85 CCTGAGAAGACATTCACCAACCCAGGAGGGTGGCTG 123

RESULT 12
US-09-599-564A-7
; Sequence 7, Application US/09599564A
; Patent No. 6362318
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahment
```

```
; APPLICANT: CHASSIN, Dorine
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 017753-127
; CURRENT APPLICATION NUMBER: US/09/599,564A
; PRIOR FILING DATE: 2000-06-23
; PRIOR FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 123
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(123)
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
US-09-599-564A-7

Query Match          28.9%; Score 99; DB 4; Length 123;
Best Local Similarity 100.0%; Pred. No. 6.5e-23;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCAGAGCTGAGGGATGTGTCCTCCGATTTGGAACACACTTGTCTCATATTGCCCATG 60
      |||||||
Db 25 GCAGAGCTGAGGGATGTGTCCTCCGATTTGGAACACACTTGTCTCATATTGCCCATG 84

Qy 61 CCTGAGAAGACATTCACCAACCCAGGAGGGTGGCTG 99
      |||||||
Db 85 CCTGAGAAGACATTCACCAACCCAGGAGGGTGGCTG 123

RESULT 13
US-09-174-465D-11
; Sequence 11, Application US/09174465D
; Patent No. 6180364
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahment
; APPLICANT: CHASSIN, Dorine
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 017753-103
; CURRENT APPLICATION NUMBER: US/09/174,465D
; PRIOR FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(93)
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
US-09-174-465D-11

Query Match          26.3%; Score 90; DB 3; Length 93;
Best Local Similarity 100.0%; Pred. No. 4.8e-20;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



QY 253 TCCCGAAAAAGAGAGTGGAGCTCACAGATTGTATCCATTCTGTTGTGAAGTAATTGT 312  
Db 1 TCCCGAAAAAGAGAGTGGAGCTCACAGATTGTATCCATTCTGTTGTGAAGTAATTGT 60  
QY 313 GACGATGGAACCTTCAGTTAAATTATGTACA 342  
Db 61 GACGATGGAACCTTCAGTTAAATTATGTACA 90

## RESULT 14

US-09-599-564A-11  
; Sequence 11, Application US/09599564A  
; Patent No. 6362318  
; GENERAL INFORMATION:  
; APPLICANT: KOMAN, Alment  
; APPLICANT: CHASSIN, Dorine  
; APPLICANT: BELLET, Dominique  
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE  
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL  
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID  
; TITLE OF INVENTION: PROTEIN  
; FILE REFERENCE: 017753-127  
; CURRENT APPLICATION NUMBER: US/09/599,564A  
; CURRENT FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: 09/174,465  
; PRIOR FILING DATE: 1998-10-19  
; PRIOR APPLICATION NUMBER: US 08/482,842  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 93  
; TYPE: DNA  
; ORGANISM: Unknown  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(93)  
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early  
; OTHER INFORMATION: Placenta Insulin-Like peptide  
US-09-599-564A-11

Query Match 26.3%; Score 90; DB 4; Length 93;  
Best Local Similarity 100.0%; Pred. No. 4.8e-20;  
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 TCCCGAAAAAGAGAGTGGAGCTCACAGATTGTATCCATTCTGTTGTGAAGTAATTGT 312  
Db 1 TCCCGAAAAAGAGAGTGGAGCTCACAGATTGTATCCATTCTGTTGTGAAGTAATTGT 60  
QY 313 GACGATGGAACCTTCAGTTAAATTATGTACA 342  
Db 61 GACGATGGAACCTTCAGTTAAATTATGTACA 90

## RESULT 15

5464756-17  
; Patent No. 5464756  
; APPLICANT: HENNER, DENNIS J.; VANDLEN, RICHARD L.; WILKINS,  
; JAMES A.; YANSURA, DANIEL G.  
; TITLE OF INVENTION: PROCESS AND COMPOSITIONS FOR THE  
; ISOLATION HUMAN RELAXIN  
; NUMBER OF SEQUENCES: 42  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/908,766  
; FILING DATE: 01-JUL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 347,550  
; FILING DATE: 04-MAY-1989  
; SEQ ID NO:17;  
; LENGTH: 489  
5464756-17

Query Match 21.0%; Score 71.8; DB 6; Length 489;

Best Local Similarity 66.5%; Pred. No. 7.9e-14;  
Matches 103; Conservative 0; Mismatches 52; Indels 0; Gaps 0;  
QY 82 ACCCCAGGAGGGTGGCTGCTGGAATCTGGACGTCCCAAGAAATGGTGTCACCTCCCAAC 141  
Db 103 AGCCAGGAGATGCTCTCTCAGACACCTAGACCAGTGGCAGAAATTTGCGCTTCCTCATC 162  
QY 142 AACAAAGATGGACAAAGCCCTTAGTAGCATCAGAAATTCATTCCTAAATTTGTCACACAGAG 201  
Db 163 AACAAAGATACAGAAACCATAAATATGATGTGCAAAATTTGTTGCTAAATTTGCCACAGGAG 222  
QY 202 CTGAAGAAACCCACTGTCTGAAGGGCAGCCATCAT 236  
Db 223 CTGAAGTTAACCCCTGTCTGAGATGCGACGCAAT 257

Search completed: October 17, 2003, 02:46:30  
Job time : 54 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2003, 02:21:42 ; Search time 183 Seconds  
(without alignments)  
4907.846 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 342

Sequence: 1 GCAGAGCTGAGGGGATGCG.....CTTCAGTTAAATTATGTACA 342

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/FCUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description                           |
|------------|-------|-------------|--------|----|---------------------------------------|
| C 1        | 69    | 20.2        | 231    | 9  | US-09-759-143-465 Sequence 465, App   |
| C 2        | 69    | 20.2        | 231    | 9  | US-09-780-669-465 Sequence 465, App   |
| C 3        | 69    | 20.2        | 231    | 9  | US-09-822-827-465 Sequence 465, App   |
| C 4        | 69    | 20.2        | 231    | 10 | US-09-895-793-465 Sequence 465, App   |
| C 5        | 69    | 20.2        | 231    | 10 | US-09-895-814-465 Sequence 465, App   |
| C 6        | 69    | 20.2        | 231    | 12 | US-10-144-678A-465 Sequence 465, App  |
| C 7        | 69    | 20.2        | 231    | 12 | US-10-294-025-465 Sequence 465, App   |
| C 8        | 69    | 20.2        | 231    | 13 | US-10-012-896-465 Sequence 465, App   |
| C 9        | 69    | 20.2        | 231    | 14 | US-10-010-940-465 Sequence 465, App   |
| C 10       | 68.6  | 20.1        | 558    | 14 | US-10-205-823-348 Sequence 348, App   |
| C 11       | 60    | 17.5        | 60     | 12 | US-09-908-975-31961 Sequence 31961, A |
| C 12       | 38.6  | 11.3        | 2574   | 9  | US-09-735-103-2 Sequence 2, Appli     |
| C 13       | 38.6  | 11.3        | 2574   | 13 | US-10-045-428A-2 Sequence 2, Appli    |
| C 14       | 35    | 10.2        | 116592 | 10 | US-09-818-512-3 Sequence 3, Appli     |
| C 15       | 35    | 10.2        | 116592 | 12 | US-10-354-065-3 Sequence 3, Appli     |
| C 16       | 33    | 9.6         | 528    | 13 | US-10-027-632-83954 Sequence 83954, A |

Sequence 298520, Sequence 111160, Sequence 6894, Appli Sequence 1, Appli Sequence 2481, App Sequence 246198, Sequence 382, App Sequence 30248, A Sequence 30249, A Sequence 30250, A Sequence 313, App Sequence 313, App Sequence 312, App Sequence 312, App Sequence 93, Appli Sequence 35, Appli Sequence 314, App Sequence 6, Appli Sequence 3396, App Sequence 23414, Sequence 134, App Sequence 3, Appli Sequence 9632, App Sequence 3869, App Sequence 185112, Sequence 187, App Sequence 215, App Sequence 8229, App Sequence 164, App

## ALIGNMENTS

RESULT 1  
US-09-759-143-465/c  
; Sequence 465, Application US/09759143  
; Patent No. US2002002248A1

## GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Steiky, Yasir A.W.  
APPLICANT: Hepler, William

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.427C23  
CURRENT APPLICATION NUMBER: US/09759,143  
CURRENT FILING DATE: 2001-01-12  
NUMBER OF SEQ ID NOS: 934  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 465  
LENGTH: 231  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-759-143-465

Query Match 20.2%; Score 69; DB 9; Length 231;  
Best Local Similarity 74.4%; Pred. No. 1.2e-12;  
Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;



US-09-895-814-465/c

; Sequence 465, Application US/09895814  
; Publication No. US20020193296A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Kalos, Michael D.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A. W.

; APPLICANT: Hepler, William T.

; APPLICANT: Henderson, Robert A.

; APPLICANT: Hural, John

; APPLICANT: McNeill, Patricia D.

; APPLICANT: Houghton, Raymond L.

; APPLICANT: Vinals de Bassols, Carlota

; APPLICANT: Foy, Teresa

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.427C26

; CURRENT APPLICATION NUMBER: US/09/895.814

; CURRENT FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 990

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 465

; LENGTH: 231

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-895-814-465

Query Match 20.2%; Score 69; DB 10; Length 231;

Best Local Similarity 74.4%; Pred. No. 1.2e-12;

Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 120 AGAATGGTGTCAACCTCCACAAAGATGGACAGCCTTAGGTACGACATCAGAATT 179

DB 137 AAAAAATGTGGCATCTTCATCAACAAAGATACAGAAACCATATAATATGATGCAGAATT 78

QY 180 CATTCTTAATTTGTCACAGAGCTGAAGACCACTGTCTGAAGGGCAGCCATCAT 236

DB 77 TGTTCCTAATTTGCCACAGGAGCTGAAGTTAACCCCTGTCTGATGATGCAGCCAGCATT 21

RESULT 6

US-10-144-678A-465/c

; Sequence 465, Application US/10144678A

; Publication No. US20030157089A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Dillon, Davin C.

; APPLICANT: Mitcham, Jennifer L.

; APPLICANT: Harlocker, Susan L.

; APPLICANT: Jiang, Yuqiu

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Fanger, Gary R.

; APPLICANT: Retter, Marc W.

; APPLICANT: Stolk, John A.

; APPLICANT: Day, Craig H.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A. W.

; APPLICANT: Hepler, William T.  
; APPLICANT: Hural, John  
; APPLICANT: McNeill, Patricia D.  
; APPLICANT: Houghton, Raymond L.  
; APPLICANT: Vinals y de Bassols, Carlota  
; APPLICANT: Foy, Teresa M.  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Deng, Ta

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.427C28

; CURRENT APPLICATION NUMBER: US/10/144.678A

; CURRENT FILING DATE: 2002-08-12

; NUMBER OF SEQ ID NOS: 1033

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 465

; LENGTH: 231

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-144-678A-465

Query Match 20.2%; Score 69; DB 12; Length 231;

Best Local Similarity 74.4%; Pred. No. 1.2e-12;

Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 120 AGAATGGTGTCAACCTCCACAAAGATGGACAGCCTTAGGTACGACATCAGAATT 179

DB 137 AAAAAATGTGGCATCTTCATCAACAAAGATACAGAAACCATATAATATGATGCAGAATT 78

QY 180 CATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCAT 236

DB 77 TGTTCCTAATTTGCCACAGGAGCTGAAGTTAACCCCTGTCTGATGATGCAGCCAGCATT 21

RESULT 7

US-10-294-025-465/c

; Sequence 465, Application US/10294025

; Publication No. US20030185830A1

; GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun

; APPLICANT: Stolk, John A.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

; FILE REFERENCE: 210121.427C29

; CURRENT APPLICATION NUMBER: US/10/294.025

; CURRENT FILING DATE: 2002-11-12

; NUMBER OF SEQ ID NOS: 1038

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 465

; LENGTH: 231

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-294-025-465

Query Match 20.2%; Score 69; DB 12; Length 231;

Best Local Similarity 74.4%; Pred. No. 1.2e-12;

Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 120 AGAATGGTGTCAACCTCCACAAAGATGGACAGCCTTAGGTACGACATCAGAATT 179

DB 137 AAAAAATGTGGCATCTTCATCAACAAAGATACAGAAACCATATAATATGATGCAGAATT 78

QY 180 CATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCAT 236

DB 77 TGTTCCTAATTTGCCACAGGAGCTGAAGTTAACCCCTGTCTGATGATGCAGCCAGCATT 21

RESULT 8

US-10-012-896-465/c

; Sequence 465, Application US/10012896

; Publication No. US20020183251A1

; GENERAL INFORMATION:



Db 292 CTGAAGGAGCCCTATCTGAGAGGCAACCATCATT 326

## RESULT 11

US-09-908-975-31961

; Sequence 31961, Application US/09908975

; Publication No. US20030165843A1

; GENERAL INFORMATION:

; APPLICANT: SHOSHAN, AVI

; APPLICANT: WASSERMAN, ALON

; APPLICANT: MINTZ, ELI

; APPLICANT: MINTZ, LIAT

; APPLICANT: FAIGLER, SIMCHON

; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLIC

; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME

; FILE REFERENCE: 36888-0005

; CURRENT APPLICATION NUMBER: US/09/908,975

; CURRENT FILING DATE: 2001-07-20

; PRIOR APPLICATION NUMBER: US 60/287,724

; PRIOR FILING DATE: 2001-05-02

; PRIOR APPLICATION NUMBER: US 60/221,607

; PRIOR FILING DATE: 2000-07-28

; NUMBER OF SEQ ID NOS: 32337

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 31961

; LENGTH: 60

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-908-975-31961

Query Match 17.5%; Score 60; DB 12; Length 60;

Best Local Similarity 100.0%; Pred. No. 5.8e-10;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 222 AGGGAGCCATCATTTGAAGAAATATCTTCCCGAAAGAGAGAGTGGACGTCACAG 281

Db 1 AGGGAGCCATCATTTGAAGAAATATCTTCCCGAAAGAGAGAGTGGACGTCACAG 60

## RESULT 12

US-09-735-103-2

; Sequence 2, Application US/09735103

; Patent No. US20010005589A1

; GENERAL INFORMATION:

; APPLICANT: Mano, Hiroyuki

; APPLICANT: Sakata, Tsuneaki

; APPLICANT: Hasegawa, Mamoru

; TITLE OF INVENTION: Promoter

; FILE REFERENCE: 50026/011002

; CURRENT APPLICATION NUMBER: US/09/735,103

; CURRENT FILING DATE: 2000-12-12

; PRIOR APPLICATION NUMBER: 09/142,529

; PRIOR FILING DATE: 1998-09-09

; PRIOR APPLICATION NUMBER: PCT/JP97/00741

; PRIOR FILING DATE: 1997-03-10

; PRIOR APPLICATION NUMBER: JP 8/54294

; PRIOR FILING DATE: 1996-03-12

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 2574

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-735-103-2

Query Match 11.3%; Score 38.6; DB 9; Length 2574;

Best Local Similarity 52.9%; Pred. No. 0.058;

Matches 83; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 54 CCCATGCGCTGAGAGACATTCACCAACCCAGAGGGTGGCTGGAATCTGGACG 113

Db 102 CCAGAAGACCGAGATGAATTTCAACACTATCTCTAGAAGAGATTTCTTATTAAGAGTCCCA 161

QY 114 TCCCAAGAAATGGTGTCAACCTCCCAACAAAGATGGACAGCCTTAGGTACGACATC 173

Db 162 GCAGAAAAGAGACATCATCTCTTAACACTACAAGAGAGACTTTGTGTACTTCCAAAATC 221

QY 174 AGAATTCATTCCTAATTTGTCCACGAGAGCTTGAAGAAA 210

Db 222 CGTGTTCAGCTACTATGAGGGTCGAGCGGAGAGAAA 258

## RESULT 13

US-10-045-428A-2

; Sequence 2, Application US/10045428A

; Publication No. US20020115845A1

; GENERAL INFORMATION:

; APPLICANT: Mano, Hiroyuki

; APPLICANT: Sakata, Tsuneaki

; APPLICANT: Hasegawa, Mamoru

; APPLICANT: Tabata, Toshiaki

; TITLE OF INVENTION: Promoter

; FILE REFERENCE: 50026/011003

; CURRENT APPLICATION NUMBER: US/10/045,428A

; CURRENT FILING DATE: 2002-04-15

; PRIOR APPLICATION NUMBER: 09/735,103

; PRIOR FILING DATE: 2000-12-12

; PRIOR APPLICATION NUMBER: 09/142,529

; PRIOR FILING DATE: 1999-08-12

; PRIOR APPLICATION NUMBER: PCT/JP97/00741

; PRIOR FILING DATE: 1997-03-10

; PRIOR APPLICATION NUMBER: JP 8/54294

; PRIOR FILING DATE: 1996-03-12

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 2574

; TYPE: DNA

; ORGANISM: Mus musculus

US-10-045-428A-2

Query Match 11.3%; Score 38.6; DB 13; Length 2574;

Best Local Similarity 52.9%; Pred. No. 0.058;

Matches 83; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 54 CCCATGCGCTGAGAGACATTCACCAACCCAGAGGGTGGCTGGAATCTGGACG 113

Db 102 CCAGAAGACCGAGATGAATTTCAACACTATCTCTAGAAGAGATTTCTTATTAAGAGTCCCA 161

QY 114 TCCCAAGAAATGGTGTCAACCTCCCAACAAAGATGGACAGCCTTAGGTACGACATC 173

Db 162 GCAGAAAAGAGACATCATCTCTTAACACTACAAGAGAGACTTTGTGTACTTCCAAAATC 221

QY 174 AGAATTCATTCCTAATTTGTCCACGAGAGCTTGAAGAAA 210

Db 222 CGTGTTCAGCTACTATGAGGGTCGAGCGGAGAGAAA 258

## RESULT 14

US-09-818-512-3/c

; Sequence 3, Application US/09818512

; Patent No. US20020142416A1

; GENERAL INFORMATION:

; APPLICANT: BEASLEY, Ellen et al.

; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL001192

; CURRENT APPLICATION NUMBER: US/09/818,512

; CURRENT FILING DATE: 2001-03-28

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 116592

; TYPE: DNA

; ORGANISM: Human

```
?
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(116592)
; OTHER INFORMATION: n = A,T,C or G
US-09-818-512-3

Query Match      10.2%; Score 35; DB 10; Length 116592;
Best Local Similarity 53.2%; Pred. No. 7;
Matches 74; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY      101 TGGAAATCTGGACGTCCTCCAAAGAAATGGTGTCACCTCCCAACACAAAGATGGACAGCCT 160
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4383 TAGCAGATGTAGGTGGCGTGGTGGCGGCGCAACATACAGCACAAGACTGAAATACCT 4324

QY      161 TAGGTACGACATCAGAAATTCCTTAATTTGTCCACGAGCTGAAGAAACCACTGTCTG 220
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4323 AATACAGGTGATTAATAATCAATTATAGTCTATTATCTGGAATAATAATAAACTGCAAT 4264

QY      221 AAGGCGAGCCCATCTTGA 239
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4263 CAGAGTAGCCAGAGCTGAA 4245

RESULT 15
US-10-354-065-3/c
; Sequence 3, Application US/10354065
; Publication No. US20030138837A1
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001192DIV
; CURRENT APPLICATION NUMBER: US/10/354,065
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 116592
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(116592)
; OTHER INFORMATION: n = A,T,C or G
US-10-354-065-3

Query Match      10.2%; Score 35; DB 12; Length 116592;
Best Local Similarity 53.2%; Pred. No. 7;
Matches 74; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY      101 TGGAAATCTGGACGTCCTCCAAAGAAATGGTGTCACCTCCCAACACAAAGATGGACAGCCT 160
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4383 TAGCAGATGTAGGTGGCGTGGTGGCGGCGCAACATACAGCACAAGACTGAAATACCT 4324

QY      161 TAGGTACGACATCAGAAATTCCTTAATTTGTCCACGAGCTGAAGAAACCACTGTCTG 220
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4323 AATACAGGTGATTAATAATCAATTATAGTCTATTATCTGGAATAATAATAAACTGCAAT 4264

QY      221 AAGGCGAGCCCATCTTGA 239
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4263 CAGAGTAGCCAGAGCTGAA 4245
```

Search completed: October 17, 2003, 03:26:36  
Job time : 187 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2003, 01:30:16 ; Search time 1428 Seconds

(without alignments)  
5820.820 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 342

Sequence: 1 GCAGAGCTGAGGGATGTGG.....CTTCAGTTAAATTAATGACA 342

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_joy:\*

19: em\_gss\_pin:\*

20: em\_gss\_vit:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

# SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| C 1        | 342   | 100.0       | 451    | 9  | AI088100    |
| 2          | 342   | 100.0       | 621    | 13 | BX102762    |
| 3          | 342   | 100.0       | 816    | 9  | AU136752    |
| 4          | 342   | 100.0       | 1201   | 9  | AL550042    |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

|      |       |      |     |    |          |
|------|-------|------|-----|----|----------|
| C 5  | 340.4 | 99.5 | 478 | 9  | AI148843 |
| C 6  | 340.4 | 99.5 | 629 | 12 | BQ011259 |
| C 7  | 331   | 96.8 | 550 | 9  | AA77897  |
| C 8  | 328.4 | 96.0 | 448 | 9  | AI128289 |
| C 9  | 321.4 | 94.0 | 523 | 9  | AI800597 |
| C 10 | 315.8 | 92.3 | 948 | 10 | BG623890 |
| C 11 | 308.4 | 90.2 | 684 | 14 | CB549181 |
| C 12 | 308.4 | 90.2 | 701 | 14 | CB548981 |
| C 13 | 307.2 | 89.8 | 424 | 12 | BQ004108 |
| C 14 | 306.8 | 89.7 | 551 | 14 | CB549885 |
| C 15 | 306.8 | 89.7 | 584 | 14 | CB549375 |
| C 16 | 306.8 | 89.7 | 646 | 14 | CB549870 |
| C 17 | 306.8 | 89.7 | 655 | 14 | CB550346 |
| C 18 | 306.8 | 89.7 | 661 | 14 | CB550055 |
| C 19 | 306.8 | 89.7 | 664 | 14 | CB549153 |
| C 20 | 306.8 | 89.7 | 665 | 14 | CB549109 |
| C 21 | 306.8 | 89.7 | 668 | 14 | CB549740 |
| C 22 | 306.8 | 89.7 | 680 | 14 | CB550616 |
| C 23 | 306.8 | 89.7 | 693 | 14 | CB549308 |
| C 24 | 306.8 | 89.7 | 726 | 14 | CB549649 |
| C 25 | 305.2 | 89.2 | 601 | 14 | CB549945 |
| C 26 | 305.2 | 89.2 | 671 | 14 | CB549798 |
| C 27 | 303.6 | 88.8 | 625 | 14 | CB549861 |
| C 28 | 297.4 | 87.0 | 569 | 14 | CB549095 |
| C 29 | 288.4 | 84.3 | 511 | 14 | CB549723 |
| C 30 | 276.6 | 80.9 | 659 | 14 | CB549812 |
| C 31 | 274.6 | 80.3 | 490 | 14 | CB549463 |
| C 32 | 265.4 | 77.6 | 414 | 9  | AI864182 |
| C 33 | 244.4 | 71.5 | 431 | 9  | AA442679 |
| C 34 | 241.2 | 70.5 | 482 | 14 | R75651   |
| C 35 | 226.4 | 66.2 | 348 | 12 | BQ007247 |
| C 36 | 212.4 | 62.1 | 427 | 14 | R62136   |
| C 37 | 211   | 61.7 | 401 | 14 | H70182   |
| C 38 | 201   | 58.8 | 329 | 14 | R68426   |
| C 39 | 199.4 | 58.3 | 414 | 14 | H02449   |
| C 40 | 191   | 55.8 | 820 | 10 | BG197868 |
| C 41 | 189.4 | 55.4 | 342 | 14 | CB549212 |
| C 42 | 183   | 53.5 | 303 | 12 | BQ007399 |
| C 43 | 181.2 | 53.0 | 705 | 14 | CB550027 |
| C 44 | 176.2 | 51.5 | 465 | 14 | R62240   |
| C 45 | 170.4 | 49.8 | 731 | 14 | CB549758 |

## ALIGNMENTS

|            |                                                                                                                                                                                                                                                                                                                                                |                                      |              |            |        |                 |
|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------|--------------|------------|--------|-----------------|
| RESULT 1   | AI088100/C                                                                                                                                                                                                                                                                                                                                     | AI088100                             | 451 bp       | mrna       | linear | EST 18-AUG-1998 |
| LOCUS      | oo20a07.x1                                                                                                                                                                                                                                                                                                                                     | Soares_NSF_F8_9W_OT_PA_P_S1          | Homo sapiens | CDNA clone |        |                 |
| DEFINITION | IMAGE:1566708 3' similar to gb:X59357 60S RIBOSOMAL PROTEIN L22 (HUMAN); mRNA sequence.                                                                                                                                                                                                                                                        |                                      |              |            |        |                 |
| ACCESSION  | AI088100                                                                                                                                                                                                                                                                                                                                       |                                      |              |            |        |                 |
| VERSION    | AI088100.1                                                                                                                                                                                                                                                                                                                                     | GI:3426808                           |              |            |        |                 |
| KEYWORDS   | EST.                                                                                                                                                                                                                                                                                                                                           |                                      |              |            |        |                 |
| SOURCE     | Homo sapiens (human)                                                                                                                                                                                                                                                                                                                           |                                      |              |            |        |                 |
| ORGANISM   | Homo sapiens                                                                                                                                                                                                                                                                                                                                   |                                      |              |            |        |                 |
| REFERENCE  | 1 (bases 1 to 451)                                                                                                                                                                                                                                                                                                                             |                                      |              |            |        |                 |
| AUTHORS    | NCI-CGAP                                                                                                                                                                                                                                                                                                                                       | http://www.ncbi.nlm.nih.gov/ncicgap. |              |            |        |                 |
| TITLE      | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index                                                                                                                                                                                                                                                              |                                      |              |            |        |                 |
| JOURNAL    | Unpublished                                                                                                                                                                                                                                                                                                                                    |                                      |              |            |        |                 |
| COMMENT    | Contact: Robert Strausberg, Ph.D.<br>Email: cgaps@email.nih.gov<br>This clone is available royalty-free through LBNL; contact the IMAGE Consortium (info@image.lbnl.gov) for further information.<br>Seq primer: -40ml3 fwd. ET from Amerham<br>High quality sequence stop: 436.<br>Location/Qualifiers<br>1. .451<br>/organism="Homo sapiens" |                                      |              |            |        |                 |



```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1566708"
/lab_host="DH10B"
/clone_lib="Soares_NSF_F8_9W_OT_PA_P1"
/notes="Organ: pooled; Vector: pT73D-pac (Pharmacia) with
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was from
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and cloneids: Soares NBHSF pool 1:
309384-310919, 323208-328695 Soares ND2HP pool 1:
145032-147335, 147320-148103, 148872-149255, 15002 -
150407, 151176-152237 Soares ND2HFR-9W pool 1:
759380-760583, 772014-774407 Soares NDHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NDHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."
Soares 95 c 97 a 144 t

```

| Query Match           | 100.0%          | Score 342;                                                     | DB 9;     | Length 451; |
|-----------------------|-----------------|----------------------------------------------------------------|-----------|-------------|
| Best Local Similarity | 100.0%;         | Pred. No. 7.7e-93;                                             |           |             |
| Matches 342;          | Conservative 0; | Mismatches 0;                                                  | Indels 0; | Gaps 0;     |
| Qy                    | 1               | CGACAGCTGAGGGGATGTGTCGCCCATTTGGAAAAACACTTCGCTGCATATTGCCCATG    | 60        |             |
| Db                    | 444             | CGACAGCTGAGGGGATGTGTCGCCCATTTGGAAAAACACTTCGCTGCATATTGCCCATG    | 385       |             |
| Qy                    | 61              | CCTGGAAGAAGATTCCACCACACCCCGAGGAGGGTGCTGCTGGAATCTGGACGTCGCCAAA  | 120       |             |
| Db                    | 384             | CCTGGAAGAAGATTCCACCACACCCCGAGGAGGGTGCTGCTGGAATCTGGACGTCGCCAAA  | 325       |             |
| Qy                    | 121             | GAATGGTGTCAACTCCACACACAAGATGGACAAGCCCTTAGTGACGATCAGAAATTC      | 180       |             |
| Db                    | 324             | GAATGGTGTCAACTCCACACACAAGATGGACAAGCCCTTAGTGACGATCAGAAATTC      | 265       |             |
| Qy                    | 181             | ATTCCTAATTTGTACACAGAGCTGAAGAAAACCACTGTCTGAAGGCGAGCCATCATTTGAAG | 240       |             |
| Db                    | 264             | ATTCCTAATTTGTACACAGAGCTGAAGAAAACCACTGTCTGAAGGCGAGCCATCATTTGAAG | 205       |             |
| Qy                    | 241             | AAATAATATCTTTCCCGCAAAAAGAGAAGTGGACGTACAGATTTTGATCCATTTCTGTTGT  | 300       |             |
| Db                    | 204             | AAATAATATCTTTCCCGCAAAAAGAGAAGTGGACGTACAGATTTTGATCCATTTCTGTTGT  | 145       |             |
| Qy                    | 301             | GAAGTAATTTGTGACGATGGAACTTCAGTTAAATTAATGTACA                    | 342       |             |
| Db                    | 144             | GAAGTAATTTGTGACGATGGAACTTCAGTTAAATTAATGTACA                    | 103       |             |

[illegible]

RZPD: IMAG998E2118.  
RZPDLIB: I.M.A.G.E. cDNA Clone Collection;  
Human Unigeneset - RZPD3 (RZPDLIB No.972)  
[http://www.rzpd.de/CloneCards/cgi-  
bin/showLib.pl.cgi?response/libNo=972](http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi?response/libNo=972) Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
[www.rzpd.de](http://www.rzpd.de)

This clone is available royalty-free from RZPD; contact RZPD ([clone@rzpd.de](mailto:clone@rzpd.de)) for further information. Seq primer: M13r, Primer sequence: TTTACACAGGAAACAGCTATGAC.

FEATURES

```

1. 021
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE998E2118 : IMAGE:66428"

```

```

/dev_stage="20 week-post conception fetus"
/lab_host="PH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen1 NF1L5"
/notes="Organ: Liver and Spleen; Vector: p7T73D (Ph
with a modified polylinker; Site1: Pac I; Site2:
last strand cDNA was primed with a Pac I - oligo(dT)
5' - RACTGGAGAGATTTATTAGATCTTTTTTTTTTTTT 3'
double-stranded cDNA was ligated to Eco RI adaptor
(Pharmacia), digested with Pac I and cloned into the
and Eco RI sites of the modified p7T73 vector. Li
went through one round of normalization. Library
constructed by Bento Soares and M. Patricia Ronaldo."

```

|            | 200 a | 141 c | 138 g | 142 t |
|------------|-------|-------|-------|-------|
| BASE COUNT |       |       |       |       |
| ORIGIN     |       |       |       |       |

| Query Match           | 100.08;         | Score 342;                                                    | DB 13;    | Length 621; |
|-----------------------|-----------------|---------------------------------------------------------------|-----------|-------------|
| Best Local Similarity | 100.08;         | Pred. No. 9,1e-93;                                            |           |             |
| Matches 342;          | Conservative 0; | Mismatches 0;                                                 | Indels 0; | Gaps 0;     |
| Qy                    | 1               | GCAGAGCTGAGGGGATGGTCCCGCATTTTGGAAAACACATTGCTGTCAATTATGCCCATG  | 60        |             |
| Db                    | 170             | GCAGAGCTGAGGGGATGGTCCCGCATTTTGGAAAACACATTGCTGTCAATTATGCCCATG  | 229       |             |
| Qy                    | 61              | CCTGGAAGACATTCACCAACCACCCAGAGGGTGGCTGTGGAATCTCGAGGTCGCCAAA    | 120       |             |
| Db                    | 230             | CCTGGAAGACATTCACCAACCACCCAGAGGGTGGCTGTGGAATCTCGAGGTCGCCAAA    | 289       |             |
| Qy                    | 121             | GAATGGTGTCACCTCCACAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC      | 180       |             |
| Db                    | 290             | GAATGGTGTCACCTCCACAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC      | 349       |             |
| Qy                    | 181             | ATTCCTAATTTGTCACAGAGCTGAGAAACCAGTGTCTGAAGGGCAGCCATCATTGAAG    | 240       |             |
| Db                    | 350             | ATTCCTAATTTGTCACAGAGCTGAGAAACCAGTGTCTGAAGGGCAGCCATCATTGAAG    | 409       |             |
| Qy                    | 241             | AAAAATAATACTTTCCCGCAAAAAGAAAGTGGAGCTCACAGATTTGATCCATCTCTGTTGT | 300       |             |
| Db                    | 410             | AAAAATAATACTTTCCCGCAAAAAGAAAGTGGAGCTCACAGATTTGATCCATCTCTGTTGT | 469       |             |
| Qy                    | 301             | GAAGTAATTTGTGACGATGGAACTTCAGTTAAATATTGTACA                    | 342       |             |
| Db                    | 470             | GAAGTAATTTGTGACGATGGAACTTCAGTTAAATATTGTACA                    | 511       |             |

|            |                                                                         |        |      |        |                 |
|------------|-------------------------------------------------------------------------|--------|------|--------|-----------------|
| RESULT 3   | AUI36752                                                                | 816 bp | mrna | linear | EST 02-AUG-2000 |
| LOCUS      | AUI36752                                                                |        |      |        |                 |
| DEFINITION | AUI36752 PLACE1 Homo sapiens CDNA clone PLACE1005004 5', mRNA sequence. |        |      |        |                 |
| ACCESSION  | AUI36752                                                                |        |      |        |                 |
| VERSION    | AUI36752.1                                                              |        |      |        |                 |
| KEYWORDS   | GI:10997291                                                             |        |      |        |                 |
| SOURCE     | Homo sapiens{human}                                                     |        |      |        |                 |

```

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 816)
JOURNAL
COMMENT Ota.T., Nishikawa.T., Suzuki.Y., Ishii.S., Saito.K., Kawai.Y.,
Yamamoto.J., Wakamatsu.A., Nakamura.Y., Nagai.T., Sugano,S. and
Isogai.T.
HRI human cDNA project
Unpublished
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
Location/Qualifiers
1..816
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1005004"
/tissue_type="placenta"
/clone_lib="PLACE1"
/note="Vector: pME185FL3"
BASE COUNT 259 a 170 c 177 g 204 t 6 others
ORIGIN
Query Match 100.0%; Score 342; DB 9; Length 816;
Best Local Similarity 100.0%; Pred. No. 1e-92;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACITTCGTGTCATATGGCCCCATG 60
Db 181 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACITTCGTGTCATATGGCCCCAIG 240
QY 61 CQTGAGAGACATTCACACACACCCAGGAGGTGGCTGGGAATCTGGAGCTGCCAAA 120
Db 241 CQTGAGAGACATTCACACACACCCAGGAGGTGGCTGGGAATCTGGAGCTGCCAAA 300
QY 121 GAAATGGTGTCAACCTCCCAACAAACAAAGATGGCAAGCCTTAGGTACGACATCAGAATTC 180
Db 301 GAAATGGTGTCAACCTCCCAACAAACAAAGATGGCAAGCCTTAGGTACGACATCAGAATTC 360
QY 181 ATTCCTTAATTTGTACACAGCTGGAAGAACACCTGTCTGAAGGCGACCCATCATTGAAG 240
Db 361 ATTCCTTAATTTGTACACAGCTGGAAGAACACCTGTCTGAAGGCGACCCATCATTGAAG 420
QY 241 AAAATAATACTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATTCCTGTGT 300
Db 421 AAAATAATACTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATTCCTGTGT 480
QY 301 GAAGTAATTTGTGACGATGGAACTTCAGTTAAATATGTACA 342
Db 481 GAAGTAATTTGTGACGATGGAACTTCAGTTAAATATGTACA 522

RESULT 4
AL550042 1201 bp mRNA linear EST 31-MAY-2003
LOCUS
DEFINITION AL550042 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0DI053YC21 5-PRIME, mRNA sequence.
ACCESSION AL550042
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 1201)
AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12886616.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5302.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/Cluster.cgi?seq=CS0DI053AB110P1&cluster=5302.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DI053AB110P1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI053YC21"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 373 a 243 c 240 g 321 t 24 others
ORIGIN
Query Match 100.0%; Score 342; DB 9; Length 1201;
Best Local Similarity 100.0%; Pred. No. 1.3e-92;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACITTCGTGTCATATGGCCCCATG 60
Db 223 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACITTCGTGTCATATGGCCCCATG 282
QY 61 CQTGAGAGACATTCACACACACCCAGGAGGTGGCTGGGAATCTGGAGCTGCCAAA 120
Db 283 CQTGAGAGACATTCACACACACCCAGGAGGTGGCTGGGAATCTGGAGCTGCCAAA 342
QY 121 GAAATGGTGTCAACCTCCCAACAAACAAAGATGGCAAGCCTTAGGTACGACATCAGAATTC 180
Db 343 GAAATGGTGTCAACCTCCCAACAAACAAAGATGGCAAGCCTTAGGTACGACATCAGAATTC 402
QY 181 ATTCCTTAATTTGTACACAGCTGGAAGAACACCTGTCTGAAGGCGACCCATCATTGAAG 240
Db 403 ATTCCTTAATTTGTACACAGCTGGAAGAACACCTGTCTGAAGGCGACCCATCATTGAAG 462
QY 241 AAAATAATACTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATTCCTGTGT 300
Db 463 AAAATAATACTTTCCCGCAAAAAGAGAGTGGAGTGCACAGATTTGATCCATTCCTGTGT 522
QY 301 GAAGTAATTTGTGACGATGGAACTTCAGTTAAATATGTACA 342
Db 523 GAAGTAATTTGTGACGATGGAACTTCAGTTAAATATGTACA 564

RESULT 5
AL148843/c 478 bp mRNA linear EST 30-SEP-1998
LOCUS
DEFINITION AL148843 x1 Soares.placenta.8to9weeks.2NDHP8to9W Homo sapiens cDNA
clone IMAGE:1714515 3' similar to gb:X59357 60S RIBOSOMAL PROTEIN
L22 (HUMAN);, mRNA sequence.
ACCESSION AL148843
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 478)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 434.
Location/Qualifiers
1..478
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1714515"
/dev_stage="two placentae; one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares.placenta_8to9weeks_2NHP8to9W"
/note="Organ: Placenta; Vector: pT73D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5',
TGTACCAATCTGAAGTGGAGCGCGCGATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."
BASE COUNT 119 a 103 c 105 g 151 t
ORIGIN
Query Match 99.5%; Score 340.4; DB 9; Length 478;
Best Local Similarity 99.7%; Pred. No. 2.4e-92;
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCAGAGCTGAGGGATGTGTCCTCCCATTTTGGAAACACTTCTGTCTATATGCCCCATG 60
DB 442 GCAGAGCTGAGGGATGTGTCCTCCCATTTTGGAAACACTTCTGTCTATATGCCCCATG 383
QY 61 CCTGAGAAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGAGCTCCCAAA 120
DB 382 CCTGAGAAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGAGCTCCCAAA 323
QY 121 GAAATGCTGTCACCTCCACACCAAGATGGACAGCCTTAGTACGACATCAGAAATTC 180
DB 322 GAAATGCTGTCACCTCCACACCAAGATGGACAGCCTTAGTACGACATCAGAAATTC 263
QY 181 ATTCCTAATTTGTCACGAGCTGAAGAAACCACTGCTGGAAGGCGAGCCATCATTTGAAG 240
DB 262 ATTCCTAATTTGTCACGAGCTGAAGAAACCACTGCTGGAAGGCGAGCCATCATTTGAAG 203
QY 241 ARAATATACTTTCCCGAAGAGAGAGTGGAGCTCAGAGTTTTCATCATTCCTGTTGT 300
DB 202 ARAATATACTTTCCCGAAGAGAGAGTGGAGCTCAGAGTTTTCATCATTCCTGTTGT 143
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 342
DB 142 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 101

RESULT 6
BQ011259/c
LOCUS
DEFINITION BQ011259 629 bp mRNA linear EST 26-MAR-2002
UI-1-BC0-afn-f-07-0-UI.s2 NCI-CGAP_P11 Homo sapiens cDNA clone
UI-1-BC0-afn-f-07-0-UI 3', mRNA sequence.
ACCESSION BQ011259
VERSION BQ011259.1 GI:19736160
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 629)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 434.
Location/Qualifiers
1..629
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-1-BC0-afn-f-07-0-UI"
/tissue_type="Placenta"
/dev_stage="8-9 weeks"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI-CGAP_P11"
/note="Organ: Placenta; Vector: pT73-Pac (Pharmacia) with
a modified polylinker; Site 1: Eco RI; Site 2: Not I;
NCI-CGAP_P11 is a cDNA library containing the following
tissue(s): Placenta 8-9 weeks pregnant. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an Eco RI
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is GA.
TAG_L18=UI-1-BC0
TAG_TISSUE=placenta human 8 week
TAG_SEQ=GA"
BASE COUNT 142 a 142 c 142 g 203 t
ORIGIN
Query Match 99.5%; Score 340.4; DB 12; Length 629;
Best Local Similarity 99.7%; Pred. No. 2.8e-92;
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCAGAGCTGAGGGATGTGTCCTCCCATTTTGGAAACACTTCTGTCTATATGCCCCATG 60
DB 455 GCAGAGCTGAGGGATGTGTCCTCCCATTTTGGAAACACTTCTGTCTATATGCCCCATG 396
QY 61 CCTGAGAAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGAGCTCCCAAA 120
DB 395 CCTGAGAAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGAGCTCCCAAA 336
QY 121 GAAATGCTGTCACCTCCACACCAAGATGGACAGCCTTAGTACGACATCAGAAATTC 180
DB 335 GAAATGCTGTCACCTCCACACCAAGATGGACAGCCTTAGTACGACATCAGAAATTC 276
QY 181 ATTCCTAATTTGTCACGAGCTGAAGAAACCACTGCTGGAAGGCGAGCCATCATTTGAAG 240
DB 275 ATTCCTAATTTGTCACGAGCTGAAGAAACCACTGCTGGAAGGCGAGCCATCATTTGAAG 216
QY 241 ARAATATACTTTCCCGAAGAGAGTGGAGCTCAGAGTTTTCATCATTCCTGTTGT 300
DB 215 ARAATATACTTTCCCGAAGAGAGTGGAGCTCAGAGTTTTCATCATTCCTGTTGT 156
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 342
DB 155 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 114

```

```

180 CATTCCTAAATTTGTCACAGAGCTGAAGAAACCACTGCTCTGAAGGGCAGCCATCATTTGAA 239
181 |
182 |
183 |
184 |
185 |
186 |
187 |
188 |
189 |
190 |
191 |
192 |
193 |
194 |
195 |
196 |
197 |
198 |
199 |
200 |
201 |
202 |
203 |
204 |
205 |
206 |
207 |
208 |
209 |
210 |
211 |
212 |
213 |
214 |
215 |
216 |
217 |
218 |
219 |
220 |
221 |
222 |
223 |
224 |
225 |
226 |
227 |
228 |
229 |
230 |
231 |
232 |
233 |
234 |
235 |
236 |
237 |
238 |
239 |
240 |
241 |
242 |
243 |
244 |
245 |
246 |
247 |
248 |
249 |
250 |
251 |
252 |
253 |
254 |
255 |
256 |
257 |
258 |
259 |
260 |
261 |
262 |
263 |
264 |
265 |
266 |
267 |
268 |
269 |
270 |
271 |
272 |
273 |
274 |
275 |
276 |
277 |
278 |
279 |
280 |
281 |
282 |
283 |
284 |
285 |
286 |
287 |
288 |
289 |
290 |
291 |
292 |
293 |
294 |
295 |
296 |
297 |
298 |
299 |
300 |
301 |
302 |
303 |
304 |
305 |
306 |
307 |
308 |
309 |
310 |
311 |
312 |
313 |
314 |
315 |
316 |
317 |
318 |
319 |
320 |
321 |
322 |
323 |
324 |
325 |
326 |
327 |
328 |
329 |
330 |
331 |
332 |
333 |
334 |
335 |
336 |
337 |
338 |
339 |
340 |
341 |
342 |
343 |
344 |
345 |
346 |
347 |
348 |
349 |
350 |
351 |
352 |
353 |
354 |
355 |
356 |
357 |
358 |
359 |
360 |
361 |
362 |
363 |
364 |
365 |
366 |
367 |
368 |
369 |
370 |
371 |
372 |
373 |
374 |
375 |
376 |
377 |
378 |
379 |
380 |
381 |
382 |
383 |
384 |
385 |
386 |
387 |
388 |
389 |
390 |
391 |
392 |
393 |
394 |
395 |
396 |
397 |
398 |
399 |
400 |
401 |
402 |
403 |
404 |
405 |
406 |
407 |
408 |
409 |
410 |
411 |
412 |
413 |
414 |
415 |
416 |
417 |
418 |
419 |
420 |
421 |
422 |
423 |
424 |
425 |
426 |
427 |
428 |
429 |
430 |
431 |
432 |
433 |
434 |
435 |
436 |
437 |
438 |
439 |
440 |
441 |
442 |
443 |
444 |
445 |
446 |
447 |
448 |
449 |
450 |
451 |
452 |
453 |
454 |
455 |
456 |
457 |
458 |
459 |
460 |
461 |
462 |
463 |
464 |
465 |
466 |
467 |
468 |
469 |
470 |
471 |
472 |
473 |
474 |
475 |
476 |
477 |
478 |
479 |
480 |
481 |
482 |
483 |
484 |
485 |
486 |
487 |
488 |
489 |
490 |
491 |
492 |
493 |
494 |
495 |
496 |
497 |
498 |
499 |
500 |
501 |
502 |
503 |
504 |
505 |
506 |
507 |
508 |
509 |
510 |
511 |
512 |
513 |
514 |
515 |
516 |
517 |
518 |
519 |
520 |
521 |
522 |
523 |
524 |
525 |
526 |
527 |
528 |
529 |
530 |
531 |
532 |
533 |
534 |
535 |
536 |
537 |
538 |
539 |
540 |
541 |
542 |
543 |
544 |
545 |
546 |
547 |
548 |
549 |
550 |
551 |
552 |
553 |
554 |
555 |
556 |
557 |
558 |
559 |
560 |
561 |
562 |
563 |
564 |
565 |
566 |
567 |
568 |
569 |
570 |
571 |
572 |
573 |
574 |
575 |
576 |
577 |
578 |
579 |
580 |
581 |
582 |
583 |
584 |
585 |
586 |
587 |
588 |
589 |
590 |
591 |
592 |
593 |
594 |
595 |
596 |
597 |
598 |
599 |
600 |
601 |
602 |
603 |
604 |
605 |
606 |
607 |
608 |
609 |
610 |
611 |
612 |
613 |
614 |
615 |
616 |
617 |
618 |
619 |
620 |
621 |
622 |
623 |
624 |
625 |
626 |
627 |
628 |
629 |
630 |
631 |
632 |
633 |
634 |
635 |
636 |
637 |
638 |
639 |
640 |
641 |
642 |
643 |
644 |
645 |
646 |
647 |
648 |
649 |
650 |
651 |
652 |
653 |
654 |
655 |
656 |
657 |
658 |
659 |
660 |
661 |
662 |
663 |
664 |
665 |
666 |
667 |
668 |
669 |
670 |
671 |
672 |
673 |
674 |
675 |
676 |
677 |
678 |
679 |
680 |
681 |
682 |
683 |
684 |
685 |
686 |
687 |
688 |
689 |
690 |
691 |
692 |
693 |
694 |
695 |
696 |
697 |
698 |
699 |
700 |
701 |
702 |
703 |
704 |
705 |
706 |
707 |
708 |
709 |
710 |
711 |
712 |
713 |
714 |
715 |
716 |
717 |
718 |
719 |
720 |
721 |
722 |
723 |
724 |
725 |
726 |
727 |
728 |
729 |
730 |
731 |
732 |
733 |
734 |
735 |
736 |
737 |
738 |
739 |
740 |
741 |
742 |
743 |
744 |
745 |
746 |
747 |
748 |
749 |
750 |
751 |
752 |
753 |
754 |
755 |
756 |
757 |
758 |
759 |
760 |
761 |
762 |
763 |
764 |
765 |
766 |
767 |
768 |
769 |
770 |
771 |
772 |
773 |
774 |
775 |
776 |
777 |
778 |
779 |
780 |
781 |
782 |
783 |
784 |
785 |
786 |
787 |
788 |
789 |
790 |
791 |
792 |
793 |
794 |
795 |
796 |
797 |
798 |
799 |
800 |
801 |
802 |
803 |
804 |
805 |
806 |
807 |
808 |
809 |
810 |
811 |
812 |
813 |
814 |
815 |
816 |
817 |
818 |
819 |
820 |
821 |
822 |
823 |
824 |
825 |
826 |
827 |
828 |
829 |
830 |
831 |
832 |
833 |
834 |
835 |
836 |
837 |
838 |
839 |
840 |
841 |
842 |
843 |
844 |
845 |
846 |
847 |
848 |
849 |
850 |
851 |
852 |
853 |
854 |
855 |
856 |
857 |
858 |
859 |
860 |
861 |
862 |
863 |
864 |
865 |
866 |
867 |
868 |
869 |
870 |
871 |
872 |
873 |
874 |
875 |
876 |
877 |
878 |
879 |
880 |
881 |
882 |
883 |
884 |
885 |
886 |
887 |
888 |
889 |
890 |
891 |
892 |
893 |
894 |
895 |
896 |
897 |
898 |
899 |
900 |
901 |
902 |
903 |
904 |
905 |
906 |
907 |
908 |
909 |
910 |
911 |
912 |
913 |
914 |
915 |
916 |
917 |
918 |
919 |
920 |
921 |
922 |
923 |
924 |
925 |
926 |
927 |
928 |
929 |
930 |
931 |
932 |
933 |
934 |
935 |
936 |
937 |
938 |
939 |
940 |
941 |
942 |
943 |
944 |
945 |
946 |
947 |
948 |
949 |
950 |
951 |
952 |
953 |
954 |
955 |
956 |
957 |
958 |
959 |
960 |
961 |
962 |
963 |
964 |
965 |
966 |
967 |
968 |
969 |
970 |
971 |
972 |
973 |
974 |
975 |
976 |
977 |
978 |
979 |
980 |
981 |
982 |
983 |
984 |
985 |
986 |
987 |
988 |
989 |
990 |
991 |
992 |
993 |
994 |
995 |
996 |
997 |
998 |
999 |

```

```

322 GAAATGGTGTCAACCTCCAAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 263
181 ATTCTTAATTTGTACACAGAGCTGAAGAAACCACTGTCTGAAGGCGACGCCATCAITGAAG 240
262 ATTCTTAATTTGTACACAGAGCTGAAGAAACCACTGTCTGAAGGCGACGCCATCAITGAAG 203
241 AAAATTAATACTTTCCCGCAAAAGAGAGTGGAGCTGCACAGATTGTGATCCATTCTGTGT 300
202 AAAATTAATACTTTCCCGCAAAAGAGAGTGGAGCTGCACAGATTGTGATCCATTCTGTGT 143
301 GAAGTAATTTGTGAGATGCGAACTTCAGTTAAATTAATATGACA 342
142 GAAGTAATTTGTGAGATGCGAACTTCAGTTAAATTAATATGACA 101

RESULT 9
AI800597/c
LOCUS      523 bp      mRNA      linear      EST 19-DEC-1999
DEFINITION w912405.x1 Soares NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
IMAGE:2364873 3' similar to SW:INL4_HUMAN Q14641 EARLY PLACENTA
INSULIN-LIKE PEPTIDE PRECURSOR ;, mRNA sequence.
ACCESSION  AI800597
VERSION     AI800597
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 523)
AUTHORS     NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE       National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            This clone is available royalty-free through LNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert length: 1083      Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 451.
FEATURES    Location/Qualifiers
             1..523
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:2364873"
             /lab_host="DH10B"
             /clone_lib="Soares NSF_F8_9W_OT_PA_P_S1"
             /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
             a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
             Equal amounts of plasmid DNA from five normalized
             libraries were mixed, and ss circles were made in vitro.
             Following HAP purification, this DNA was used as tracer in
             a subtractive hybridization reaction. The driver was
             PCR-amplified cDNAs from pools of 5,000 clones made from
             the same 5 libraries. The pools consisted of the following
             libraries and clones: Soares NBHSF pool 1:
             309384-310919, 323208-325895 Soares NB2HP pool 1:
             145032-147335, 147720-148103, 148872-149255, 15002 -
             150407, 151176-152327 Soares NB2HPF8-9W pool 1:
             758280-760583, 772104-774407 Soares NBHPA pool 1:
             304776-306311, 320136-322823, 326280-326663 Soares NBHQT
             pool 1: 723720-728407, 739080-740999 Subtraction by Bento
             Soares and M. Fatima Bonaldo."
BASE COUNT  133 a 111 c 117 g 162 t
ORIGIN
Query Match      94.08; Score 321.4; DB 9; Length 523;
Best Local Similarity 98.08; Pred. No. 1.5e-86;
Matches 336; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

1 GCAGAGCTGAGGGATGTGTCGCCCGGATTTGGAAACACTTGTGTCATATTGCCCATG 60

```

```

. Db      448 GCAGAGCTGAGGGATGTGTCGCCCGGATTTGGAAACACTTGTGTCATATTGCCCATG 389
Qy      61  CCTGGAACACATTCACACACACCCAGGAGGCTGCTGGAATCTGGAGCTGCCAAA 120
Db      388  CCTGGAACACATTCACACACACCCAGGAGGCTGCTGGAATCTGGAGCTGCCAAA 329
Qy      121  GAAATGGTGTCAACCTCCAAACAAGATGGACAAGCCTTAGGTACGACATC-AGAAAT 179
Db      328  GAAATGGTGTCAACCTCCAAACAAGATGGACAAGCCTTAGGTACGACATC-AGAAAT 269
Qy      180  CATTCTTAATTTGTACACAGAGCTGAAGAAACCACTGTCTGAAGGCGACGCCATCAATGAA 239
Db      268  CATTCTTAATTTGTACACAGAGCTGAAGAAACCACTGTCTGAAGGCGACGCCATCAATGAA 209
Qy      240  GAAATTAATACTTTCCCGCAAAAGAGAGTGGAGCTGCACAGATTGTGATCCATTCTGTG 299
Db      208  AAAATTAATACTTTCCCGCAAAAGAGAGTGGAGCTGCACAGATTGTGATCCATTCTGTG 149
Qy      300  TGAAGTAATTTGTGAGATGCGAACTTCAGTTAAATTAATATGACA 342
Db      148  TGAAGTAATTTGTGAGATGCGAACTTCAGTTAAATTAATATGACA 106

RESULT 10
BG623890      948 bp      mRNA      linear      EST 18-APR-2001
LOCUS      602649042F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4770395 5',
DEFINITION mRNA sequence.
ACCESSION  BG623890
VERSION     BG623890
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 948)
AUTHORS     NIH-MGC http://mgi.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LUCM1637 row: d column: 12
            High quality sequence stop: 624.
FEATURES    Location/Qualifiers
             1..948
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /clone="IMAGE:4770395"
             /lab_host="NIH_MGC_79"
             /clone_lib="NIH_MGC_79"
             /note="Organ: placenta; Vector: pDNR-LIB (Clontech);
             Site_1: SfiI (ggccattatggcc); Site_2: SfiI (ggccattatggcc)
             ); 5' and 3' adaptors were used in cloning as follows: 5'
             adaptor sequence: 5'-CACGGCCATTTATGCCC-3' and 3' adaptor
             sequence: 5'-ATTCTAGAGCGGAGCGCGGACATG-dt(30)BN-3'
             (where B = A, C, G, or T). Average
             insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies
             contained inserts by PCR. This library was enriched for
             full-length clones and was constructed by Clontech
             Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
             Library."
BASE COUNT  329 a 201 c 254 g 164 t
ORIGIN

```

Query Match 92.3%; Score 315.8; DB 10; Length 948;  
 Best Local Similarity 98.8%; Pred. No. 1e-84;  
 Matches 339; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 GCAGAGCTGAGGGGATGTGCTGCCCGATTTGGAAACACTTGTCTCATATTGCCCCATG 60  
 DB 200 GCAGAGCTGAGGGGATGTGCTGCCCGATTTGGAAACACTTGTCTCATATTGCCCCATG 259

QY 61 CTTGAGAGACATTCACACACCCAGAGGGTGGCTGCTGGAATCTGGAGCTGCCAAA 120  
 DB 260 CTTGAGAGACATTCACACACCCAGAGGGTGGCTGCTGGAATCTGGAGCTGCCAAA 318

QY 121 GAA-ATGCTGCTCAACTCCCAACAACAAGATGACAAAGCCTTAGGTACCATCAGAAAT 179  
 DB 319 GAAATGCTGCTCAACTCCCAACAACAAGATGACAAAGCCTTAGGTACCATCAGAAAT 378

QY 180 CATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCCATCATTTGAA 239  
 DB 379 CATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCCATCATTTGAA 438

QY 240 GAAATTAATTAATTTGCCGCAAAAGAGAGTGCACATCTTGCATTTGATCCATCTGTTG 299  
 DB 439 GAAATTAATTAATTTGCCGCAAAAGAGAGTGCACATCTTGCATTTGATCCATCTGTTG 498

QY 300 TGAAGTAAATTTGTCAGATGAAGAACTTCAGTTAAATTAATGTACA 342  
 DB 499 TGAAGTAAATTTGTCAGATGAAGAACTTCAGTTAAATTAATGTACA 541

RESULT 11  
 CB549181  
 LOCUS  
 DEFINITION MPMPL0003\_F01 MPMPL Macaca mulatta cDNA, mRNA sequence.  
 ACCESSION CB549181  
 VERSION CB549181.1 GI:31298376  
 KEYWORDS EST.  
 SOURCE Macaca mulatta (rhesus monkey)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopithecinae; Macaca.

REFERENCE 1 (bases 1 to 684)  
 Katze, M.G., Bumgarner, R., Korth, M., Feldman, R., Amjadi, M. and Holzman, T.  
 Expressed sequence tags from Rhesus macaque placenta  
 Unpublished  
 Contact: Holzman T  
 Katze Lab  
 University of Washington  
 Box 358070, Seattle, WA 98195-8070, USA  
 Tel: 206 732 6156  
 Fax: 206 732 6055  
 Email: ted@locke.hs.washington.edu  
 Similar to GenBank entry BC026254 Homo sapiens, insulin-like 4 (placenta), clone MGC:22499 IMAGE:4770395, mRNA, complete cds. 4/2002  
 Plate: MPMPL0003 row: F column: 01.

FEATURES  
 source  
 1..684  
 /organism="Macaca mulatta"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9544"  
 /sex="male"  
 /dev\_stage="neonatal"  
 /clone\_lib="MMPL"  
 /note="Organ: placenta"  
 215 a 147 c 150 g 172 t

BASE COUNT 215 a 147 c 150 g 172 t  
 ORIGIN  
 Query Match 90.2%; Score 308.4; DB 14; Length 684;  
 Best Local Similarity 93.9%; Pred. No. 1.5e-82;  
 Matches 321; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGCTGCCCGATTTGGAAACACTTGTCTCATATTGCCCCATG 60  
 DB 146 GCAGAGTTGAAGGGATGTGCTGCCCGATTTGSAAGAGGATGCTGTCTCATATTGCCCCATG 205

QY 61 CTTGAGAGACATTCACACACCCAGAGGGTGGCTGCTGGAATCTGGAGCTGCCAAA 120  
 DB 206 CTTGAGAGACATTCACACACCCAGAGGGTGGCTGCTGGAATCTGGAGCTGCCAAA 265

QY 121 GAAATGCTGCTCAACTCCCAACAACAAGATGACAAAGCCTTAGGTACCATCAGAAATTC 180  
 DB 266 GAAACGGTGTCAACTCCCAACAACAAGATGACAAAGCCTTAGGTACCATCAGAAATTC 325

QY 181 ATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCCATCATTTGAG 240  
 DB 326 ATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCCATCATTTGAG 385

QY 241 AAAATAATTAATTTGCCGCAAAAGAGAGTGCACATCTTGCATTTGATCCATCTGTTG 300  
 DB 386 AAAATAATTAATTTGCCGCAAAAGAGAGTGCACATCTTGCATTTGATCCATCTGTTG 445

QY 301 GAAGTAAATTTGTCAGATGAAGAACTTCAGTTAAATTAATGTACA 342  
 DB 446 GAAGTAAATTTGTCAGATGAAGAACTTCAGTTAAATTAATGTACA 487

RESULT 12  
 CB548981  
 LOCUS  
 DEFINITION MPMPL0017\_E12 MPMPL Macaca mulatta cDNA, mRNA sequence.  
 ACCESSION CB548981  
 VERSION CB548981.1 GI:31298176  
 KEYWORDS EST.  
 SOURCE Macaca mulatta (rhesus monkey)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 Cercopithecinae; Macaca.

REFERENCE 1 (bases 1 to 701)  
 Katze, M.G., Bumgarner, R., Korth, M., Feldman, R., Amjadi, M. and Holzman, T.  
 Expressed sequence tags from Rhesus macaque placenta  
 Unpublished  
 Contact: Holzman T  
 Katze Lab  
 University of Washington  
 Box 358070, Seattle, WA 98195-8070, USA  
 Tel: 206 732 6156  
 Fax: 206 732 6055  
 Email: ted@locke.hs.washington.edu  
 Similar to GenBank entry BC026254 BC026254 Homo sapiens, insulin-like 4 (placenta), clone MGC:22499 IMAGE:4770395, mRNA, complete cds. 4/2002  
 Plate: MPMPL0017 row: E column: 12.

FEATURES  
 Location/Qualifiers  
 1..701  
 /organism="Macaca mulatta"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9544"  
 /sex="male"  
 /dev\_stage="neonatal"  
 /clone\_lib="MMPL"  
 /note="Organ: placenta"  
 221 a 145 c 156 g 179 t

BASE COUNT 221 a 145 c 156 g 179 t  
 ORIGIN  
 Query Match 90.2%; Score 308.4; DB 14; Length 701;  
 Best Local Similarity 93.9%; Pred. No. 1.6e-82;  
 Matches 321; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGCTGCCCGATTTGGAAACACTTGTCTCATATTGCCCCATG 60  
 DB 146 GCAGAGTTGAAGGGATGTGCTGCCCGATTTGSAAGAGGATGCTGTCTCATATTGCCCCATG 205